DOCUMENT RESUME

ED 059 627	EM 009 668

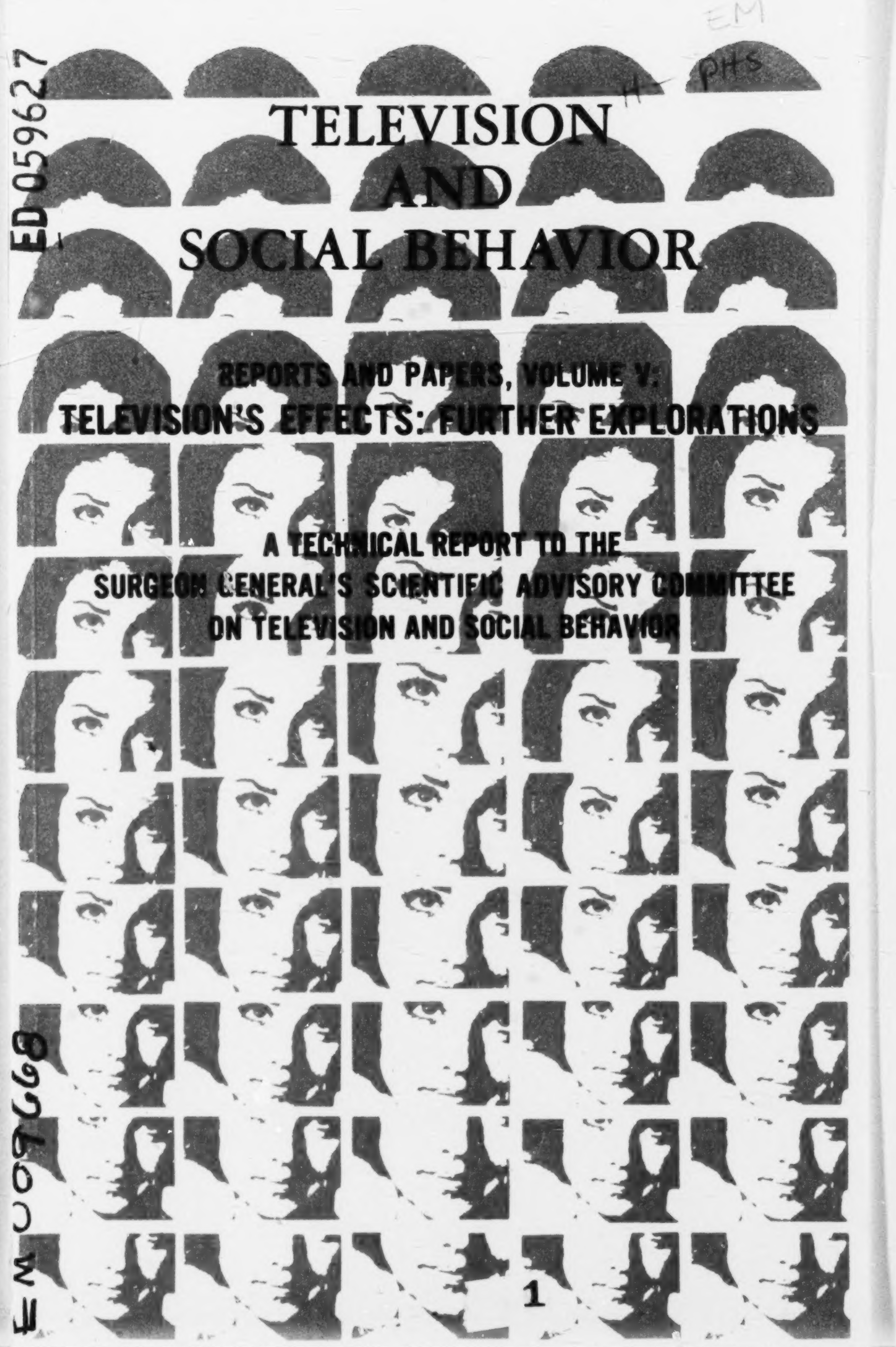
AUTHOR	Comstock, George A., Ed.; And Others
TITLE	Television and Social Behavior; Reports and Papers, Volume V: Television's Effects: Further Explorations.
INSTITUTION	Surgeon General's Scientific Advisory Committee on Television and Social Behavior, Washington, D.C.
PUB DATE	Apr 72
NOTE	383p.; See also EM 009 435, EM 009 664, EM 009 665, EM 009 666, EM 009 667
AVAILABLE FROM	Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (DHEW Pub. No. HSM 72-9060)
EDRS PRICE	MF-$0.65 HC-$13.16
DESCRIPTORS	*Aggression; Arousal Patterns; *Commercial Television; Comprehension; Emotional Experience; Emotional Response; Identification (Psychological); Negro Attitudes; Perception; Self Concept; Social Behavior; *Television Research; Television Viewing; *Violence

ABSTRACT

The studies in this fifth volume of technical reports to the Scientific Advisory Committee on Television and Social Behavior explore the role of mediated aggression and violence on relevant human behaviors in some new and interesting ways. The volume begins with an overview of the studies, presenting a statement of the principal questions asked, an overview of the main variables and methods, a summary of each study, and conclusions. The major research areas dealt with by the different investigators include facial expressions as indicators of emotional reactions while viewing televised violence, dream content as an index of certain viewer reactions to televised violence, familiarity with violence as a factor in predisposing the viewer to see violence more readily, emotional arousal and the mediation of the effects of communication messages, learning from messages manipulated in terms of violent content and color television, and black consciousness and identification in relation to perception of violence. (Author/SH)

ED 059627

EM
H PHS



TELEVISION AND SOCIAL BEHAVIOR

REPORTS AND PAPERS, VOLUME V:
TELEVISION'S EFFECTS: FURTHER EXPLORATIONS

A TECHNICAL REPORT TO THE
SURGEON GENERAL'S SCIENTIFIC ADVISORY COMMITTEE
ON TELEVISION AND SOCIAL BEHAVIOR

EM 009668

ED 059627

TELEVISION
AND
SOCIAL BEHAVIOR

REPORTS AND PAPERS, VOLUME V:
TELEVISION'S EFFECTS: FURTHER EXPLORATIONS

A TECHNICAL REPORT TO THE
SURGEON GENERAL'S SCIENTIFIC ADVISORY COMMITTEE
ON TELEVISION AND SOCIAL BEHAVIOR

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Edited By
George A. Comstock, Eli A. Rubinstein, and John P. Murray
Editorial Coordination: Susan Lloyd-Jones

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Health Services and Mental Health Administration

National Institute of Mental Health
5600 Fishers Lane
Rockville, Maryland

Staff Members

Eli A. Rubinstein	Vice Chairman, Surgeon General's Scientific Advisory Committee
George A. Comstock	Senior Research Coordinator
John P. Murray	Research Coordinator
Michael Adler	Staff Assistant
Eileen Marchak	Research Assistant
Susan S. Lloyd-Jones	Editor
Joseph D. Reckley	Administrative Officer
Margaret D. Salladay	Secretary
Laura A. De Lisi	Secretary

Former Staff Members

Douglas A. Fuchs	Senior Research Coordinator (through 6/70)
John P. Robinson	Research Coordinator (through 9/70)
Harold Leigh	Administrative Officer (through 10/70)
Thomas Brubeck	Information Officer (through 5/71)
Deborah Cutler	Research Assistant (through 8/70)
Jan W. Lipkin	Secretary (through 4/70)

Advisory Committee Members

Ira H. Cisin
Thomas E. Coffin
Irving L. Janis
Joseph T. Klapper
Harold Mendelsohn
Eveline Omwake
Charles A. Pinderhughes
Ithiel de Sola Pool
Alberta E. Siegel
Anthony F. C. Wallace
Andrew S. Watson
Gerhart D. Wiebe

Preface

This document is one of five volumes of technical reports resulting from a broad scientific inquiry about television and its impact on the viewer. In the spring of 1969, by Congressional request, the DHEW initiated a special program under the general auspices of a Surgeon General's Scientific Advisory Committee on Television and Social Behavior. The major emphasis was to be on an examination of the relationship between televised violence and the attitudes and behavior of children. During the ensuing two years, more than fifty scientists participated directly in this program of research and produced over forty scientific reports.

The reports which are included in these five volumes are the independent work of the participating researchers. These results have all been made available to the Scientific Advisory Committee as evidence which the Committee could then evaluate and draw its own conclusions in the preparation of its own report. However, this work is of significance in its own right and is being published independently as source material for other researchers and for such interest as the general public may have in these technical reports.

In any broad scientific undertaking of this nature, where many individuals are involved, a careful balance between collaboration and independence of responsibility must be established. During the two and half years that this program of research was active, a constant effort was made to protect the scientific independence of the individual investigators and, at the same time: 1) to foster both cooperation and exchange among the researchers, 2) to develop as much of a total program structure as possible, and 3) to permit maximum communication and feedback among the researchers, the full-time staff responsible for planning and implementing the total research program, and the Scientific Advisory Committee responsible for the final assessment and evaluation of the research.

This is not the place to describe in detail how that balance of collaboration and independence was established and maintained. I believe, however, that these five volumes of technical reports provide an accurate and meaningful indication of our success in achieving the goal. The reports themselves are the products of the respective authors. They have been edited only to insure some comparability of format and to delete any excessive redundancies in review of the literature or introductory material. In some instances, where a report seemed initially too long the author was requested to reduce the report without deleting any critical material. All editing done by staff was submitted for the author's approval. We believe the result has made each of these five volumes a more readable and integrated totality than would otherwise be expected from a collection of research reports produced under the time constraints of this program.

In each instance, the integration of the five volumes was further established by the inclusion of an overview paper which attempts to summarize and relate the papers in the volume. These overview papers are also the independent work of the respective authors.

It would be difficult to convey to the reader the extraordinary efforts required by all participants in this research program to bring the endeavor to its published conclusion within the time allotted. Despite that time pressure, these volumes demonstrate an unusually high level of both productivity and quality for an area of research which has had more than its share of complexity and controversy.

In addition to the work of all persons directlly engaged in this program, a very large number of individuals at one time or another provided advice and guidance to the researchers, to the staff, and to the scientific Advisory Committee. It would be impossible to provide a complete list of these additional consultants. The total count is in the hundreds. While their names are not visible in these products, their counsel was often a very significant factor in the course of an individual piece of research or in a decision on the direction of the research program. To all those individuals, this program owes a special debt of gratitude for the collective wisdom made available to us.

And finally, on behalf both of the members of the Scientific Advisory Committee and of the staff who served the program, I wish especially to express much appreciation to the participating researchers who did the work and wrote the reports that contributed the new knowledge contained in these volumes.

Eli A. Rubinstein
Vice-Chairman, Surgeon General's
Scientific Advisory Committee on
Television and Social Behavior

Contents

Page

Televised Violence: Further Explorations (Overview).
Bradley S. Greenberg 1

Facial Expressions of Emotion White Watching Televised Violence as Predictors of Subsequent Aggression.
Paul Ekman, Robert M. Liebert, Wallace V. Friesen, Randall Harrison, Carl Zlatchin, Edward J. Malmstrom, and Robert A. Baron 22

Televised Violence and Dream Content.
David Foulkes, Edward Belvedere, and Terry Brubaker 59

Race, Identification, and Television Violence.
Cedric Clark 120

Social Class and Racial Differences in Children's Perceptions of Television Violence.
Bradley S. Greenberg and Thomas F. Gordon 185

Children's Perceptions of Television Violence: A Replication.
Bradley S. Greenberg and Thomas F. Gordon 211

Children's Violence Perception as a Function of Television Violence.
Martin S. Rabinovitch, Malcolm S. McLean, Jr., James W. Markham, and Albert D. Talbott 231

Violence and Color Television: What Children of Different Ages Learn.
Natan I. Katzman 253

Studies in Film- and Television-Mediated Arousal and Aggression: A Progress Report.
Percy H. Tannenbaum 309

Television and Aggression: A Discussion

Catharsis of Aggression Among Institutionalized Boys: Fact or Artifact?
Robert M. Liebert, Michael D. Sobol, and Emily S. Davidson ... 351

Television and Aggression: A Reply to Liebert, Sobol and Davidson.
Seymour Feshbach and Robert D. Singer 359

Catharsis of Aggression Among Institutionalized Boys: Further Discussion.
Robert M. Liebert, Emily S. Davidson, and Michael P. Sobol ... 366

Television and Aggression: Some Reactions to the Liebert, Davidson and Sobol and Review and Response.
Seymour Feshbach and *Robert D. Singer* 373

Televised Violence: Further Explorations (Overview)

Bradley S. Greenberg

The studies in this volume explore the role of mediated aggression and violence on relevant human behaviors in some new and interesting ways. The studies provide alternative conceptualizations of the problem area, the addition of some first-time variables, or striking changes in the conditions of observing behavior. In all cases, they are not intended to be definitive; most authors urgently caution that their immediate report is an introductory one, that the research effort on which it is based is continuing, that the data presented are undergoing additional analyses, that followup studies are in progress, and so on. Despite these cautions, the reports are contained in this volume, and this introductory paper attempts to cope with them.

This coping poses certain issues. For one, the writer must encompass his own research program. One's judgments are not unaffected by biases, both theoretical and methodological. For a second, the notion of an integrative paper would be faulty. The issues are too diverse, the measures too varied, the subject populations too amorphous to allow integration of all which follows.

This paper presents: (a) a statement of the principal questions raised by the several investigators; (b) an overview of their main variables and methods; (c) a summary of each of the studies which identifies what we judge to be the major findings and limitations; and (d) a concluding section.

Although these tasks may seem overly safe, it is not the present assignment to make policy recommendations or to thoroughly evaluate each project. Rather, we have attempted to extract from the several hundreds of pages of text, tables, and appendices those information units of most use to the Surgeon General's Scientific Advisory Committee in preparing its own final report.

The emphasis, description, summary, and implications attributed to these several projects remain the sole responsibility of this writer. Given the proximity of the full reports, the reader is invited to make his own comparisons where so inclined.

THE MAJOR RESEARCH QUESTIONS

Ekman et al. asked:

(a) Does facial information obtained while watching televised violence yield reliable measures of emotional reactions?

(b) Do such immediate emotional reactions as assessed from facial behaviors relate to postviewing aggressive behavior?

Foulkes et al. asked:

(a) Does viewing violence just before going to sleep affect the intensity or vividness of recalled dream content, its pleasantness, its aggressive content, or its guilt and anxiety components?

Greenberg and Gordon[1] asked:

(a) Is the perception of violence—its amount, its acceptability, its realism, and its enjoyment—influenced by the young viewer's economic status and/or race?

Katzman asked:

(a) Is more information learned (peripheral material in particular) from a program high in violent content?

(b) Is the visual component of such learning enhanced by the color presentation of a program high in violent content?

Rabinovitch et al. asked:

(a) Does exposure to televised violence predispose the viewer to perceive violence when both violence and nonviolence are simultaneously available to perceive?

(b) Does exposure to televised violence predispose the viewer to more readily and correctly report perceiving a violent act when separate violent and nonviolent acts are available to perceive?

Tannenbaum asked:

(a) To what extent can aggressive responses to televised violence be more generally explained as intensified behavior induced by emotionally and physiologically arousing content exposure?

(b) What is the interaction between such a general arousal formulation and cognitive factors specifically associated with violent content?

Clark asked:

(a) Will the race, and the racial consciousness, of a viewer of a television drama influence the way in which he identifies with the characters in the drama? How will the race and racial consciousness of the fictional characters affect the viewer's identification with them?

Thus, all the investigators exposed some viewers to material that was operationalized as high or low in violent content. Some used other types of content as well; still others used no content controls. Some looked at the viewers' perceptions of the content or their ability to perceive the content; some at perceptions of the viewers' looking at the content; still others at postviewing behaviors—both asleep and waking. Most asked questions in addition to those cited, but these seemingly represent the major energy foci of the various experiments.

These projects used a variety of film or television stimuli, across different age groups, some with different sexes, each in its own data-gathering setting. Numbers of subjects varied widely. Table 1 summarizes this descriptive information. The specific use of the filmed stimuli, in terms of the manipulations or variables they represented, and the dependent behavioral measures will be examined in the project summary section.

For all, straightforward exposure to one or another of a set of content manipulations was followed by assessment of some dependent behavior.

PROJECT SUMMARIES

Here, we will describe in somewhat greater detail what each investigator examined, his independent and dependent measures, and the major results.

Facial Expressions of Emotion While Watching Televised Violence as a Predictor of Subsequent Aggression *(Ekman et al.)*

Facial expressions as indicators of emotional reactions were examined here. The investigators posited that facial behavior while watching

Table 1: Methods overview

Project director	Content, length of stimulus	Number, gender of Ss	Age	Study setting
Ekman	Scenes from The Untouchables vs. sports (3½ minutes each) vs. two commercials (one minute each).	30 boys 35 girls	5-6 years old	Room at Fels Research Institute
Foulkes	Two episodes of a children's western (NBC). One with 26 elements of violence, one with three. About 50 minutes.	40 boys	10-12 years old	Sleep laboratory
Greenberg	Eight violent scenes (four/subject) from seven prime time shows, ranging from 15 seconds to 37 seconds. Different kinds of violence. Two nonviolent scenes.	588 boys	263 from 8th, 325 from 5th grades	School rooms
Katzman	One episode of Felony Squad in original form vs. same episode with most of the violence edited out (20 minutes each).	240 boys	80 each from 4th, 6th, and 9th grade	School rooms
Rabinovitch	One episode of Peter Gunn vs. one episode of Green Acres. Each 20-23 minutes.	24 girls 33 boys	6th grade	School library and cafeteria
Tannenbaum	An erotic film, violent scenes from Body and Soul, (mainly boxing) neutral travel film, humor clips, Norman McLaren film, shootout scenes, plus numerous variations of these. Time varied from 90 seconds to 5-6 minutes.	About 20 Ss per cell in each study	University students	Experimental lab on campus

televised violence would reflect emotional reactions. Further, they expected that certain of these reactions would be positively associated with subsequent aggressiveness. The contention was that the specific nature of the emotional arousal would determine the particular type of subsequent behavior.

More specifically, child viewers whose faces expressed happiness or interest while watching mayhem were expected to be more aggressive than children emoting disgust, sadness, or fear

A videotape recording was made of each child's face while watching an appropriate television stimulus. The recording made during two 25-second episodes of violence and nonviolence plus 30 seconds while watching two commercials were coded by groups of judges (college undergraduates). The videotapes were subdivided into smaller samples so that no judge made more than one assessment of a single subject. All subjects were a portion of the group studied by Liebert and Baron in another project in this NIMH study program.

Faces were coded in terms of 11 affective emotions on nine-point scales. Four were bipolar: pleasant-unpleasant; interested-disinterested; aroused-unaroused; and involved-uninvolved. Seven were single anchor scales: anger, happiness, disgust, fear, pain, sadness, and surprise. Coding was done for each subject for his three exposure samples.

Two aggressive situations were presented after exposure to the TV program. In one, the subject could press either a HELP or a HURT button, presumably facilitating or hampering a same-sex child playing a game in an adjacent room. Total length of pressing the buttons over 20 trials further helped or hurt. A second measure was the child's play behavior, as rated by an observer.

The major findings were supportive of the theoretic rationale—for boys only. To the extent that the judged emotions connoted pleasantness or happiness, a significant correlation was obtained with hurting in the postviewing situation; sadness yielded a significant negative correlation and happiness a significant positive correlation with aggressive play. These correlations were in the range of .5–.6, which is rather striking, given the small number of subjects (N=15). No consistent correlations between emotions and aggressiveness existed for the nonviolent exposure group or for exposure to commercials. Multiple correlations of these emotions with aggressiveness ranged from .71–.87 in the violence exposure group.

For the helping dependent measure, unpleasantness, pain, disinterest, lack of involvement, and lack of arousal were all positively associated. This was further support of the basic ideas studied.

All these findings applied only to the male subjects. No consistent correlations were found among the girls, although the girls showed a very similar range of emotions and a similar range in levels of postviewing aggressiveness.

Analyses of variance supported the correlational data. Among the boys viewing violence, those showing the most pleasantness, the least sadness, and the most interest performed significantly more hurting behavior or were faster to hurt.

Comparing the extremes of emotion between those exposed to violence and those exposed to nonviolence, additional support was found for these same three emotions. For example, those showing the most pleasantness while viewing violence were more aggressive than those showing most pleasantness while viewing nonviolence. Thus, there was support for the specificity of the emotional reaction in anticipating the specific later behavior.

In this last set of analyses, contradictory findings were found among the girls. Those girls who viewed violence and displayed the least pleasantness and the most sadness were more aggressive than girls with parallel emotions while watching nonviolence.

Discussion. Most imperative would be to see the comparable analyses among the older age group included in the Liebert and Baron experiment. Given the limited number of subjects here, there is need for replication with more subjects and perhaps with a more diverse set of subjects, in terms of social class or race for example. Further variations in content within the class of violent content appear warranted—e.g., relative emotional reactions to verbal aggression as well as physical aggression; violence with depicted consequences to victim and/or aggressor as well as no consequences; and ways in which the facial reactions relate to more chronic attitudes toward aggression?

The main bind is the lack of findings among the girls. Ekman et al. posit three main explanations: the lack of female actors in violent episodes, the depiction of male forms of violence, and differential female socialization with regard to attitudes toward aggression.[2]

Televised Violence and Dream Content *(Foulkes (et al.)*

This research effort sought to determine the usefulness of dream content as an index of certain viewer reactions to televised violence. The interest in the use of dream behavior was based on the following premises, several of which have prior research support:

a. Long-range and immediate effects should be sought at a level where latent and symbolic fantasy processes are operative.
b. Short-run changes in waking drive states yield expressions of related content in dreams.
c. Dream content reflects chronic personality characteristics.
d. Dream content reveals affective and cognitive reactions in a symbolic and relatively involuntary level of free responses, thus having certain advantages over more intrusive response measures.

Research questions were posed, rather than directional hypotheses. Mainly, the study examined whether presleep viewing of a violent program, in contrast to a nonviolent one, affected (reduced or increased) the vividness, pleasantness, aggression and anxiety components of dream content. This was determined for the principal variable of violent/nonviolent content, for other manipulated variables like the degree of subject involvement in the exposure situation, and for one nonmanipulated variable—the subject's normal level of exposure to television violence.

The lack of stipulated predictions stems from the contradictory findings of two prior studies in this research program, both directed by Foulkes. In one, 24 subjects in their twenties viewed a violent and nonviolent episode of a television western. Described as generally uninvolved in the task, their dream recall after watching violence was more exciting, more interesting, and more imaginative without an influence on their dreams' aggressive content or pleasantness. Thus, there was increased general activation.

The second study alternately exposed 32 children, 6-8 and 10-12 years old, to a violent western and a baseball documentary. Exposure to the western yielded dream content that was less imaginative and with less goodness of recall. The children were more involved in the violent episode. Here, there was decreased general activation.

The present study maintained the principal independent variable of violence and controlled for two factors which were thought to contribute to the contradictory findings of the earlier studies—degree of involvement and general level of exposure to television violence. The latter was measured by regularity of viewing some 18 locally shown violent programs. The former was manipulated in a manner described below.

Measures of dream content included two indices of intensity: seven of the dream's hedonic quality or pleasantness, nine of hostile content, and seven of defensiveness. The attributes were rated by two judges from a typescript prepared from a tape recording of subject interviews.

All subjects participated in an adaptation night so that they would become accustomed to the subsequent procedures and requests to be made of them. At experimental sessions—two nights seven to nine days apart—groups of two or three subjects came to the Sleep Laboratory, changed into bed-clothes, had electrodes attached to their face and scalp, were exposed to one stimulus, and slept. The procedure was to awaken each subject after ten minutes of a rapid-eye-movement sleep period, and to do so for the first four such periods that occurred. The subject was interviewed, then went back to sleep after each awakening.

Under a condition of focal involvement, their bedroom was not lit save for the television screen; the door was closed, the volume was relatively loud, and there were no interruptions. Under incidental involvement, there was bright lighting, an open door, noise from outside, three deliberate interruptions by an experimenter, and available diversions to watching television like magazines or puzzles.

Results showed that for the 25 dream content variables, not one was significantly related to the film manipulation. Four were related to the involvement manipulation (with focal involvement generally eliciting more hostile dream content, irrespective of film treatment), and one was related to prior exposure. The few significant interactions, all including involvement, did not have a consistent pattern. We quote Foulkes et al.:

> The findings of the present study indicate that the violent film did not have any systematic effect on dream hostility, anxiety, guilt, hedonic tone, or overall vividness and intensity. They thus are in direct agreement with the results of neither of the earlier two studies relating media violence to dream content.

Whereas the first study showed an increase in general activation, and the second study showed a decrease, the present study showed no difference. And it uncovered no differences on any of the variables studied here for the first time.

It should be noted, however, that in none of the three studies was there any evidence that the aggressive films increased sleep disturbance or hostility, anxiety, and negative feeling tone in reported dreams.

Discussion. When faced with three different sets of findings from three different experiments in a continuing research program, and when the most recent test is the most precise and carefully executed, one is tempted to give up on dreams as a path to the understanding of media effects on aggressiveness.

Yet we pose two precautionary notions. Several projects described in this paper determined the extent to which the experimental stimulus was familiar to the subjects. In no other study did more than a handful of subjects acknowledge familiarity. In this project, nine of 20 high exposure Ss had seen or thought they had seen the violent stimulus, four of 20 in the low exposure subgroup; seven in each exposure group had seen the nonviolent stimulus. Surely, this may be more contaminating than the use of a more original or novel stimulus, given that re-exposure to an old television program may trigger a variety of memories, both correct and not. Dream recall after exposure to new and old programming may be a heuristic question in itself.

Second, most, if not all, of the dependent dream content variables were measures based on judgments of verbalized recall samples from the subjects. The premise is that individual differences in verbal ability are randomized across experimental treatments. Yet, data are presented which negate a portion of that premise. High exposure subjects differed from low exposure ones in one significant way: they were generally lower in social class. Indeed, other research demonstrates that such children have lesser verbal ability. Certainly in terms of the kinds of verbalizations that upper-middle-class experimenters and judges would code for vividness, etc., some check on potential individual verbal abilities would appear warranted in subsequent research.

Social Class and Racial Differences in Children's Perceptions of Television Violence *(Greenberg and Gordon)*

The world of violence is make-believe, *vis a vis* movies and television, for most Americans. But in some sectors, real-world violence, hostility, and aggression are more prevalent. This is particularly so where economic conditions are poorest, and thus it is even more so for black Americans than for white. This study postulated that for disadvantaged youngsters, greater frequency of exposure to violence at home and in the neighborhood has a systematic impact on their perceptions of televised violence. The specific direction of that influence was expected to be one of sating or diminishment of the low-income youngster's perceptions of the severity of televised violent acts.

Television vignettes were taped from prime time entertainment programs. Eight different scenes of violence were used, with each viewer receiving four plus two nonviolent scenes. Each set contained an instance of violence against property, nonfatal violence against another person, a killing by gunfire, and a second killing by other means.

Perceptions of all the acts viewed were obtained on scaled-response items. In a first study, the items tapped four attitudinal dimensions: (1) amount of perceived violence; (2) acceptability of the viewed behavior; (3) liking for each of the scenes; and (4) judged reality of the viewed behavior. In the second study, a fifth percept, judged humor, emerged.

Fifth-grade boys were subjects in the first study, eighth graders in the replication. They were clustered in these independent variable categories: low-income blacks, low-income whites, middle-income whites, and upper-middle-income whites. Testing was done in groups of 5-8 in a room in their public school. Each scene—practice, experimental, and control—was rated immediately after viewing.

Figure 1 presents the basic results of the two studies.

FIGURE 1: Summary of findings

Dimensions of judgment	Study 1 (5th grade)	Study 2 (8th grade)
1. Perceived violence	Blacks perceive less violence	No difference
2. Perceived acceptability	Lower-income and blacks find it more acceptable	Blacks find it more acceptable
3. Professed liking	Lower income and blacks like it more	Lower-income and blacks like it more
4. Perceived humor	(not assessed)	Lower-income find it more humorous
5. Perceived reality	Lower-income see it as more realistic	Lower-income and blacks see it as more realistic

The findings for the younger children were more consistent for the individual scenes of violence and the collective set.

In both studies, there was evidence that a response set to television content in general could not account for the findings. Response patterns to the control scenes were quite different than to the violence scenes. This was more evident among the fifth graders.

A subhypothesis that weapon-induced violence could be perceived as more violent was supported each time. A second subhypothesis that weaponless scenes would be judged as more realistic was supported only among the younger children.

Discussion. One would have preferred to have a followup results section, which began, ". . .and the results of such differential perceptions of televised violence had these consequences in terms of postviewing behaviors. . ." But that phase of the research has not yet been completed.

The authors proposed that the racial differences noted were extensions of economic conditions, rather than cultural influences. That remains to be specifically tested. Further, perceived violence cannot be differentiated with either independent variable among the older boys. Several other findings are weaker for that group. Some form of adaptation to violence may occur with maturation—or is it increased habituation to televised violence?

All subgroups said that the violent scenes contained a substantial degree of violence. The differences were of degree, not on opposite sides of a scale's midpoint.

One wishes to begin to determine the factors leading to particular perceptions or judgments. When the low-income child says he likes violence more, what criteria does he apply? Ekman et al. believe their female data suffered because of lack of female television characters; all the characters in these scenes of violence were white. The implications of violence with opposite and mixed racial characters merits further exploration.

Finally, these data would be even more meaningful if comparable data were obtained from these same children about their perceptions of real-life acts of violence. Does what is obtained here pertain to all violence, or primarily to fantasy violence?

Children's Violence Perception As a Function of Television Violence *(Rabinovitch et al.)*

This study reasoned that familiarity with violence was a factor which predisposed a viewer to see it more readily. This would be heightened, the argument continues, in two kinds of ambiguous situations—where violence and nonviolence are equally available to perceive, and where there is an opportunity to label a stimulus, which does or does not have violent content, as having it.

Familiarity with violence was operationalized as exposure to a violent program followed by immediate testing of perceptions. A comparable group of subjects watched a nonviolent program.

Perceptions were assessed with two measures. (1) Nine pairs of slides were presented tachistoscopically (.5 seconds). Each pair contained one violent act and one nonviolent act, and all pairs were presented twice to control for eye dominance. Free response descriptions of the content were obtained for each of the 18 presentations. (2) Two hundred slides (100 violent and 100 not) were presented singly for .02 seconds each. For each slide, the subject first stated whether there was or was not hurting or fighting, and then stipulated whether he was "very sure" or "not very sure" about the judgment made.

Contradictory results were found. With the paired-slides measure, the subjects who had just seen a violent program were least likely of the three groups (Violence, No Violence, No Program) to indicate having perceived a slide depicting violence. With the individual slides measure, those most recently exposed to television violence reported seeing more acts of hurting or fighting and were significantly more likely to be more sure of their perceptions. With the paired slides, in fact, the No Program group perceived the largest number of violent acts; with the individual slides, the No Violence group was least likely to say a given slide had hurting in it and were least certain of their assessment when they did report some hurting.

Discussion. Rabinovitch et al. posed three central issues in attempting to account for this anomaly. We shall supplement these and add two of our own. Clearly, the instructions to the subjects for the second measure—and it was always the second measure—were more closely linked cognitively to the violent stimulus. Those instructions used hurting or fighting as a referent, possibly making such responses more socially acceptable for those who had been exposed to the violent program and/or cuing them more directly to the violence. In parallel fashion, watching the violent program may have raised inhibitions against seeing (or, more correctly, against reporting seeing) one violent act in a pair of acts. From the study design, one cannot separate the possible order effect of the two measures from the possible decrement in inhibition over time.

Our own addendum is that the perception of violence in the paired slides measure was physically more difficult than seeing no violence; i.e., its occurrence across all groups was far less than chance would have predicted. Given 18 opportunities to see violence, no study group mean exceeded 3.0, and one subgroup did not even see an average of one violent incident among 18 shown. Rabinovitch provided us with data on the individual slide pairs. For six of 18 pairs, two or fewer of 57 total subjects reported the violent slide. For only three pairs did 10 or more subjects see the violent member of the pair.

Finally, a modified study design is necessary to test simultaneously for the possible inhibition and propensity to see violence notions. In one condition, both members of the slide pair should contain violence. Then, if there are differential amounts of nonviolence perceived, one could meaningfully attribute the results to an inhibition mechanism. A second condition with nonviolence in both pairs would further serve to assess the affinity of those most familiar with violence to perceive it—especially when there is none to perceive.

Studies in Film- and Television-Mediated Arousal and Aggression *(Tannenbaum)*

Tannenbaum's report on his contract research to date is a tentative one. Inasmuch as his contract period extends for a full year beyond the date of the current report, neither his statement nor this summary should be considered as more than a "here's where we are today" memorandum.

The research effort is based on an intensive examination of the role of emotional arousal in mediating the effects of communication messages. In its near-original form, the theoretical model was reasoned thus:

a. Many communication messages elicit different degrees of generalized emotional arousal.
b. In particular, such arousal is likely to be evoked by messages which feature stimuli often found in violent or erotic content.
c. The generalized arousal which is elicited has drive potential.
d. The drive potential should serve to increase the degree of subsequent behavior an individual exhibits.

Therefore, heightened arousal, perhaps irrespective of its content source, should lead to more responsiveness, whether the response called for is aggression, or helping, or something else.

This basic formulation has been modified subsequently to take cognitive factors into account. Tannenbaum acknowledges that cognitive message cues, like the aggression in violent content, may interact with generalized arousal to contribute to even higher levels of aggression than arousal itself might produce. Or the aggressive cues may act independently to produce higher subsequent aggressiveness.

Such a modification implies that an individual who undergoes a relatively high level of generalized arousal during exposure to violence may use information obtained from the content to label his state of excitation. Thus, he would be more likely to "behave in a manner compatible with that label."

The basic experimental paradigm has been consistent throughout Tannenbaum's studies. First, there is an encounter between a subject and an experimenter (or his confederate), in which the subject is angered. This has been done in two ways: (1) where testing is with one subject at a time, the anger is induced through a series of mild electric shocks; (2) in group testing, the subjects are grossly insulted by a substitute teacher. After such a precondition, subjects receive a filmed stimulus for an appropriate experimental condition. After exposure, response situations are established for the subject. Where aggressiveness is the response condition, parallel measures for single and group testing have been created. For the former, the subject administers electric shocks, both in number and intensity. For the latter, subjects have an opportunity to rate the teacher in terms of whether the school system should make an offer of a regular teaching position. Where rewarding is the response condition, subjects can give tokens, worth 1-10 points, which later may be exchanged for cash.

Arousal has been assessed by a variety of physiological measures, including GSR (a composite index of sympathetic activation) and respiration.

Before summarizing the major experimental tests of the arousal theory, let us report three studies which examined the effects of the precondition of angering the subjects:

(1) Whether angering occurred before or after the exposure condition made no difference in amount of subsequent aggression.

(2) Whether pre-exposure angering was moderate or severe made no difference on a set of physiological arousal measures obtained during exposure.

(3) Angering resulted in significantly more aggression against the specific angering agent than alternative targets offered to the subjects. In a second response situation in the same experiment, subjects were given an opportunity to aggress against their original tormentor under three conditions: their shocking of him hadn't worked in round 1; his shocking by someone else should have been performed by this subject; or no one had tested him before. For all, aggressiveness increased from round 1 to round 2. This counters notions of drive reduction, and supports the target characteristic approach—i.e., that the subject's principal goal is to punish the initial aggressor.

The major experimental findings of this progress report are categorized in terms of content manipulations, response variations, and arousal patterns.

Variations in message content. 1. When an erotic film (highest in arousal), an aggressive film, and a neutral film (lowest in arousal) were used, subjects exposed to the erotic, low aggression film were significantly more aggressive than those exposed to the high aggression stimulus. Aggressiveness to the neutral film was significantly lower than either of the other two film treatments.

2. When an erotic film with no sound track, and the same film with separate nonaggressive and aggressive audio inputs were used, there was more aggression in the erotic plus aggressive stimulus.

3. When an arousing, but not aggressive, humorous film, an aggressive one, and a neutral one were used, the humorous film induced significantly more aggression than the neutral film but significantly less than the prizefight. So the cognitive properties of violence may have accounted for the latter finding, or the humor stimulus may have been less arousing.

4. Holding level of aggression relatively constant, and creating three degrees of arousal, the most and least arousing stimuli were significantly different in terms of subsequent aggression.

Variations in response. 1. Subjects rated the humor in a film clip after exposure to an aggression stimulus, a humorous one (different from the one rated), and a neutral film. Higher ratings were given the first two conditions than the last, with the first two not being different from each other.

2. Subjects rewarded the confederate (who had earlier angered, helped, or not interacted with the subject) after exposure to either the erotic or neutral stimulus. Rewards were greater for the high arousal main effect; rewards were minimum for the angering precondition main effect. A significant interaction between precondition and arousal (maximum reward with high arousal and pre-exposure helping) offers support for both the arousal-drive model and the cognitive cue proposition.

3. Both reward (for correct answers) and punishment (for wrong ones) were available to subjects who had either been in a "negative encounter" precondition, receiving eight shocks in 10 trials, or a "positive encounter," with two shocks. All subjects then were exposed to the boxing film. Significantly more aggression followed the negative encounter condition, while the reward behavior in both situations was essentially the same. This re-emphasizes the utility of including some cognitive notions in the arousal-drive formulation.

Patterns of arousal. 1. A happy ending to the prizefight stimulus was created and compared with two no-ending versions which differed in length. Less aggression followed the happy ending version, but the difference was not statistically significant.

2. Three versions of the same stimulus produced patterns of declining arousal (after the story's climax), rising arousal (just short of the climax), and early arousal (as arousal begins to build). Reward and punishment responses were available. Punishments in the declining arousal condition were significantly less than the other treatments. No differences were obtained in reward patterns. Given that the only precondition was one of angering, selectivity of effect—i.e., punishment differences only—supports the cognitive-arousal theory intersect.

3. A pilot study in which 18 subjects were exposed to essentially the same short sequence of violence on three consecutive days indicated no significant changes in physiological arousal from repeated exposures. Tannenbaum reported three kinds of response patterns. For some subjects, there was a steady increase in arousal, for others a steady decrease, and for still others an up-down pattern.

Discussion. Arousal enhances subsequent behavior. The cognitive compatibility of both the pre-exposure condition and the stimulus with the postviewing condition enhances behavior. The precise nature of the interrelationship between these remains the basic issue.

Whatever turns you on is not an inappropriate concern here. Given the abundance of television violence and the considerable fascination of large numbers of viewers with such material, the propensity is for violent content to be the principal available source of arousal.

Controversy continues as to whether a pre-angering condition is necessary or merely sufficient. Whether nontormentor targets evoke more aggression than no pre-anger condition remains to be determined.

Finally, we would urge this line of research be extended to a population of younger viewers—preschoolers through young adolescents. Although the theories examined should be independent of an age or maturation attribute, the individual differences in arousal which Tannenbaum finds have implications for less sophisticated subjects than university students. If youngsters are more easily aroused, or aroused to a greater degree by the same stimuli, the subsequent behaviors should be more intense.

Violence and Color Television: What Children of Different Ages Learn *(Katzman)*

This study differed from all others in the Television and Social Behavior program in that the dependent behavior of sole interest was information gain. There was no concern with attitudes toward aggression, nor with aggressiveness, but only with straightforward learning from messages manipulated in terms of content and format.

The rationale for the project was simple: With more attention, there should be more learning. Violent acts are an attention-getting device in television presentations, principally because of their high and focused action content; therefore, there should be more learning from a message high in violence than from a very comparable message low in violent content. The investigator differentiated peripheral from central learning, the latter being content which was coder-judged as having "to do with the plot or main action of the story, or concerns any important part of the program," and the former as having "little to do with the plot or action, or nothing to do with the main activity." He posited that violent content, while contributing to an overall learning effect, would have its principal influence in the learning of peripheral material.

A second independent variable was color, such that television stimuli were presented in both black and white and color. Prior research cited indicated that color's main influence would be in the form of additional learning of peripheral-visual material. Color was conceived of as an additional visual cue, gaining the viewer's attention and thereby stimulating learning in the manner indicated for violent content. The interaction hypothesis followed that violent content in color would result in even better recall of peripheral-visual material.

Other independent variables were age of the viewer (fourth, sixth, and ninth-grade subgroups were used) and both immediate and delayed (two weeks) recall tests, but hypotheses for these variables were not related to the television content manipulations.

The stimulus consisted of two edited versions of the same television program. In the 20-minute program, 2-3 minutes differed. In the high violence version, 2-3 minutes of nonviolent content were removed. Left in were a "struggle, a slap, a glass thrown against a wall, a point-blank murder, a gun battle, and a violent fistfight." In the low violence version, these acts were cut: the murder was implied; a chase was seen, but not the gun battle, and only the final blow of the fight.

The recall (learning) measures were developed so that items to be learned were equally available in both the high violent and low violent episodes. Thus, no items dealt with recall of violent incidents. Fifteen central learning items were used, eight dealing with auditory material and seven with visual; the peripheral items had the same total and subgroup distribution. Subjects also were asked to order 12 still photographs, showing scenes common to both television versions, as a second measure of central-visual learning. Through combining subsets of these measures, taking difference scores, etc., a total of 11 different (but not completely independent) recall indices comprised the dependent variables.

The results of the study showed no support for the hypotheses about high *vs.* low violence. Indeed, the only significant result was for more peripheral learning from the low violence exposure than from the high violence one. The peripheral-central difference, particularly in terms of visual components, also strongly favored the low violence exposure condition.

The main effects of the color-black/white manipulation supported Katzman's reasoning that color would facilitate peripheral-visual learning relative to central visual material, but this was attributable mainly to the greater learning of central visual material in the black/white treatment group. For our interests, the postulated color-violence interaction was also significant; with high violence and color, there was significantly better recall of peripheral visual material relative to central visual than in the parallel black/white treatment.

Of descriptive interest was an unanticipated consistency of interactions between violence level and grade. For six of the learning measures, the fourth graders learned more from the low violence program, the ninth graders learned more from the high violence, and the sixth graders learned equivalently from both. These paralleled the respective ratings of liking for the programs.

Discussion. Katzman proposed that for the younger children, the violence—liked less—may have reduced attentiveness to program elements, rather than stimulating it. This would account for their lesser learning, but this requires direct testing.

In retrospect, one would like to have seen a measure of viewer aggressiveness after exposure to these treatments. The relationship between the learning of violent content and its potential for modeling or imitation (e.g., the social learning discussed in another volume of these reports) is of immediate interest. Does violence in color, with or without more recall, influence subsequent behavior? More so among those who like it more or less? More so for younger or older children? No assessment was made of the learning of the violent content because it could be learned in only one condition. There is reason to believe that it is exactly that specific learning which might predict postviewing aggressiveness.

Subsequent research, which posits attentiveness as a mediating mechanism for greater learning, must directly assess such attention. If there was not differential attention between the film conditions, then the lack of theoretical support is not surprising. A classroom situation for testing may contribute to a ceiling on attentiveness.

Finally, we again miss the girls. Do they learn the same things? As well? One muses that most projects lack female staff members, and that the remainder anticipate the type of anomalous findings obtained by Ekman et al. when females are indeed added to the study dimensions.

Race, Identification, and Television Violence (*Clark*)[3]

In a study designed to measure identification with characters in a television drama, 71 teenagers (38 white and 33 black) recruited through local newspapers participated in two sessions. In the first, they filled out a background questionnaire. In the second—about four weeks later—they viewed a one-half-hour tape of a program from the *Dragnet* television series. The program featured a black militant, a black policeman, and a white policeman.

The subjects viewed the program under one of four experimental conditions: (1) racially mixed groups in a white locality; (2) racially unmixed groups in a white locality; (3) racially mixed groups in a black locality; and (4) racially unmixed groups in a black locality. Preliminary analysis showed that the locality factor had little effect on the subjects, so the groups were combined for analysis, leaving a race of subject by racial composition of group (2 x 2) design.

Dependent measures. Identification was measured in a variety of ways, which were grouped into three categories.

Three measures of identification were derived from social learning theory. In these measures, the subject/viewers were asked to answer: (1) true-false questions about the nonverbal behavior of each of the main characters; (2) open-ended questions requiring knowledge about the general behavior of the main characters; (3) questions asking them to attribute quotations from the program to the correct speaker and listener among the main characters. The investigator hypothesized that identification with any character, as measured by these three attention measures, would be highest where the race of the viewer was identical with the race of the character.

Three measures were based on social psychological theory of identification. In these measures, the viewers were asked (a) to rate the characters' actions for their "goodness" and "successfulness." They were also asked (b) to rate the degree to which each action was attributable to internal forces (to the character's own attributes and powers) or (c) to external forces (to forces outside the character's control). These latter two measures were based on the theory that "an individual tends to maintain a self-image by attributing his failures or incompetencies to his environment (externally) and his successes to his dispositions or character (internally) . . . .To the extent that a person identifies with another, he will attribute behavioral causality to that person in the same way as he would to himself." Again, Clark hypothesized that identification would be greatest where the race of the viewer and of the character were identical.

Three measures of identification were derived from Freudian theory. One asked the viewer to rate each character's "friendliness"; another asked for a rating of the character's "competency." A third, which asked for a rating of the character's "aggressiveness," was intended to measure the degree to which the viewers "identified with the aggressor" in the filmed drama.

Results. Few differences were found between black and white subjects. On the learning measures, white subjects scored consistently higher than blacks in recall of verbal material; no differences between racial groups were found in recall of nonverbal content. Both black and white subjects recalled more correct information about the two policemen than they did about the black militant. White subjects had higher mean scores on open-ended measures which asked subjects to supply specific information about things in the program. Racially mixed viewing groups had higher mean scores than racially unmixed groups.

For the social psychological measures of identification, the most clearcut finding was that blacks identified with the white policeman more than did the whites. Blacks and whites in racially mixed viewing sessions tended to identify more with the black policeman. There was a general tendency, consistent across a variety of characters and character actions, for whites to attribute causality more externally, and for blacks to attribute it more internally.

On the Freudian measures, black subjects rated the two policemen as more friendly than did white subjects. Blacks showed a slight tendency to evaluate the black militant as more friendly than did whites, but the tendency was not significant. Competency of the characters was perceived similarly by both black and white viewers, and all subjects under all conditions rated the black militant as more aggressive than the other characters.

Black consciousness and identification. The investigator hypothesized that viewers with high black consciousness would identify more with the black militant in the program than would the low black consciousness viewers, and that the low consciousness viewers would identify more with the policemen. To examine this hypothesis, the sample of black subjects was divided on the basis of a median split for their black consciousness scores.

The only significant difference found was that high black consciousness subjects tended to attribute the bad actions of the black militant to his environment rather than to personality, while low consciousness subjects did not.

CONCLUDING REMARKS

Direct comparisons from study to study are not possible in this set of projects. They approach different questions in different ways, rather than the same question by the same or different means.

Perhaps the most promising exercise, however, would be to attempt to sequence some of these notions which appear to have a related pattern to them, and to note regularities and inconsistencies.

What is and is not labeled violent has been delineated most frequently by content analysts like Gerbner (whose report appears in Volume 1 of this series). That the expert analyst may differ in degree from the ordinary viewer is apparent. That subgroups of viewers may differ is evident from the Greenberg data. The amount of violence viewed, the acceptability of the perceived behavior, the liking for that kind of thing, etc., are more pronounced among some viewers than others. These data indicate an economic background distinction as a major factor, race and perhaps age as lesser components.

We find that where such a difference in perceptions does exist, at least in terms of liking for or pleasantness of violence content, facial reactions (Ekman et al.) may be used as a less intrusive measure. Exposure to such material is at least as likely to induce more ready recognition of other instances of violence, as a portion of the Rabinovitch et al. project attests. Such recognition may more readily occur among those who favor it most. Further, and most important, if that perceived violence is well-liked, then differential liking signals the probability of more aggressive behavior (Ekman et al.). The reverse also holds: emotional signals of dislike, such as sadness, are predictive of less subsequent aggression. Thus, there is this initial patterning of relative attitudes toward violence, the labeling of emotions accompanying certain of the attitudes, and some dimensions of later behavior.

This reasoning is supported by waking behavior: there is only ambivalence as to its manifestation in dream states (Foulkes). This *post hoc* set of linkages would be even stronger had Katzman examined and obtained a relationship between learning of violence and subsequent aggressiveness, particularly if one learns more of what one enjoys more, as Katzman's data do attest.

Ekman's minimal data pool is amply supplemented by Tannenbaum, who consistently demonstrates that more aggressiveness occurs after exposure to aggression than after exposure to nonviolence or no program at all. But that was not Tannenbaum's basic objective, and here he and Ekman (and Liebért) to some extent part theoretical ways. Ekman, like Berkowitz, attributes the increased aggression to the specific cues of the violent content, in interaction with specific emotional arousal states. Ekman made one test for generalized arousal effects and obtained none in terms of frequency or duration of aggressiveness. But the bulk of the Tannenbaum evidence, across studies, across stimuli, and across subject pools, provides strong support for some generalized arousal impact.

Where Tannenbaum gets the strongest arousal, he gets the highest level of aggressiveness. That this may interact with the cognitive cues available in violent content *per se*, his most current conceptualization acknowledges. Is arousal more important than and different from aggressive content cues in eliciting aggression? What is the nature of their independent or interactive contributions to aggressiveness? These are exciting and important theoretical distinctions. But the argument does no damage to this most salient point: More overt aggressive behavior follows exposure to violent content than to nonviolent content or no content in these experimental settings.

That point has been made before. It continues to be made and is now supplemented by these explorations into factors which accompany such behavior.

FOOTNOTES

1. Only the perception-of-violence studies from the Greenberg project are reported here. Other studies are described elsewhere in this series.

2. Another report bearing on this latter issue is Dominick, J.R., and Greenberg, B.S., "Attitudes toward violence: the interaction of television exposure, family attitudes, and social class," in Volume 3 of this series.
3. The complete Clark report was not available to Professor Greenberg at the time this overview was prepared. This section of the paper was written by the staff of the Television and Social Behavior program and does not include a discussion section. The Clark paper is not treated in the summary section of this overview.

Facial Expressions of Emotion While Watching Televised Violence as Predictors of Subsequent Aggression

Paul Ekman, Robert M. Liebert, Wallace V. Friesen, Randall Harrison, Carl Zlatchin, Edward J. Malmstrom, and Robert A. Baron

Although the effects of viewing televised violence on children's subsequent aggressive behavior have been investigated in many studies, only limited progress has been made in understanding the circumstances under which such effects are most and least likely to occur. Previous approaches to this problem have generally used one of two techniques: (a) varying the type of violence, the victim, the aggressor, and/or the consequences; and (b) varying viewer characteristics (sex, age, social behavior, and so on).

In the present experiment, we employed a new approach: considering the viewer's *emotional* responses to the televised violence.[1] More specifically, this experiment was designed to determine: (a) if it is feasible to derive measures of emotion from the facial behavior of television viewers; and (b) if differences among viewers' immediate emotional reactions to television violence are related to subsequent aggressive behavior—if children who look happy when someone is killed on television, for example, will subsequently behave more aggressively than those who do not.

Research on the face and emotion. Advances during the past decade in both method and theory show that it is possible to derive valid and reliable measures of emotional response from recordings of facial behavior. Facial behavior in response to television should be particularly informative when the TV fare does indeed have some emotional impact, and when the viewer is unaware that his facial behavior is being recorded. While this expectation rests on a strong foundation of theory and research, some of the key supporting evidence is only now reaching the research literature. A brief review of this research tradition may, therefore, be helpful.

In their extensive review of more than a half-century's scientific research on the face and emotion, Ekman, Friesen, and Ellsworth (1971) conclude (a) that people can interpret, with high accuracy, facial expressions of emotion, and (b) that past studies which appear to contradict this conclusion are demonstrably flawed with methodological problems and should be reevaluated. During the past ten years, Ekman and his associates have found consistent support for this conclusion in studies using still and motion picture records (photos, films, and videotape), subjects in such diverse literate cultures as the United States, Japan, Brazil, Argentina, and Chile, and subjects in two preliterate cultures, the Fore and the Dani of New Guinea (Ekman, Sorenson, and Friesen, 1969; Ekman and Friesen, 1971).

In the Japanese-American study (Ekman, in press), adult males in the two cultures were shown stressful and nonstressful motion pictures. People in each culture then judged videotape recordings made of the facial behavior shown during film watching. In this study, both Japanese and American subjects were able to distinguish facial expressions during stress-watching from expressions during nonstress-watching, when they observed members of their own culture *and* members of the other cultures. The findings demonstrate that untrained observers can distinguish pleasant from unpleasant emotional expressions even on faces from another culture and another racial physiognomy.

As part of their work establishing cross-cultural universals in facial expression, Ekman and Friesen (1971) further found that untrained subjects, even in preliterate cultures, can reliably discriminate not only between pleasant and unpleasant affect, but among a set of specific emotions like happiness, sadness, anger, fear, surprise, and disgust. In short, the ability to judge specific, facially expressed emotions appears to be a robust phenomenon even in cultures with little exposure to Western man or to modern mass media.

In current studies, the ability to judge emotion has been found to extend—with interesting variations—to facial expressions seen very briefly (for 1/125th of a second), to subjects who are severely disturbed (as in schizophrenia or depression), and to subjects under the influence of marijuana or alcohol. All of this research points to the potential of using judgments made by untrained subjects as a well-calibrated instrument for measuring the emotions revealed in facial expression.

A second, more powerful methodology for measuring emotion from facial expression has also been developed during the last ten years. Ekman, Friesen, and Tomkins (1971) have developed the Facial Affect Scoring Technique (FAST) which employs sets of trained coders to score muscle movements in three areas of the face. They have found strong evidence for the validity of the FAST methodology in identifying the emotions conveyed by facial photographs. Similarly, Ekman (in press) found in the Japanese-American study that FAST was effective in coding motion as well as still pictures and that the resulting discriminations were more powerful and precise than those gathered from untrained observers.

While there was good reason to believe that facial measures could provide valuable data in the present study, this research extends previous work on the face and emotion in three important ways: (1) It places facial measures in a predictive rather than a postdictive frame. That is, while previous studies have typically asked whether an antecedent condition or stimulus can be inferred from facial expression, this study asks whether important social consequences can be predicted from emotional states as revealed in facial expression. (2) Facial expressions are elicited by everyday stimulation like television fare rather than by stress films. (3) The expressors are children rather than adults, females as well as males.

In designing this experiment, the investigators planned to use both observers' judgments and FAST scoring and to compare their effectiveness. Only the observers' judgments have been obtained to date. The findings from this approach do provide an answer to the primary substantive question: whether facial expressions of emotion during the viewing of televised violence are related to subsequent aggressive behavior.

Emotion and television violence. The idea that measuring the facial expressions of the viewer would be useful was based on the assumption that viewers who see the same violent episode on television may often have diverse emotional reactions. We further hypothesized that these different emotional reactions would be related to the incidence and frequency of subsequent aggression. Most simply stated: viewers who seem *happy* and *interested* when viewing violence might be expected subsequently to engage in more aggressive behavior than those who seem *unhappy, sad, disgusted, pained, afraid,* or *disinterested.*

One basis for this expectation is derived from Tomkins's (1962, 1963) theory of emotion, which suggests that experiencing positive affect is intrinsically rewarding, while experiencing negative affect is intrinsically punishing. If the viewer has a rewarding experience when viewing violence, he is apt to engage in that behavior himself; if his experience is punishing, he is not likely to engage in that behavior.

Another basis for the same expectation is the idea that positive emotional reactions when viewing a violent episode might be an index of some form of identification with the aggressor, while negative emotional reactions might be an index of some form of identification with the victim. Those who identify with the aggressor might be more likely than those who identify with the victim to engage subsequently in behavior like the aggressor's.

These formulations are illustrative of the reasoning which could lead to the expectation that specific emotional reactions of the viewer would be related to his or her subsequent aggressive behavior. However, it was not the aim of this study to investigate the source of differences among viewers' emotional reactions to televised violence, whether the source be expectations, mood, attitude, personality, peer relationships, school, family background, or whatever. That is a later question which is important only if the feasibility study demonstrates that emotional reactions when viewing televised violence do indeed predict subsequent aggression.

The present experiment was conducted with a subsample of the children also studied by Liebert and Baron (1971). Boys and girls first viewed part of an actual television program depicting either violence or nonviolent sports; they were then given an opportunity to engage in aggressive behavior against a peer. Videotape records were taken of the children's facial expressions of emotion while they watched television. Liebert and Baron examined the relationship between the two types of television inputs and subsequent aggression, and found that, over all, those children who viewed violence later engaged in more aggressive behavior than those who viewed the sports program. The present experiment examined whether the children's emotional reactions while watching the violent program predicted the incidence of subsequent aggressive behavior.

Our general hypothesis was that measurements of facial expression of emotion shown during the viewing of television violence are related to measurements of subsequent aggressive behavior. More specifically, we predicted that emotional expressions of *happiness, pleasantness,* and *interest* would be positively related to subsequent aggression, while expressions of *unpleasantness, sadness, fear, disgust, pain,* and *disinterest* would be negatively related to subsequent aggression. No prediction was made about expressions of *surprise, involvement,* or *arousal,* which could occur with either positive or negative feelings. No prediction was made about expressions of *anger*, because the measurement procedure did not allow precise determination of when the angry expression occurred. If shown in phase with the aggressor's attack, anger would be a positive predictor, but if it were out of phase with the aggressor's activity or in response to the victim's suffering, it might well be negatively related to subsequent aggression.

METHOD

The subjects were five- and six-year olds, 30 boys and 35 girls. The sample included all of the subjects of this age group studied by Liebert and Baron.[2] Each child was brought to Fels Research Institute by a parent in response to a newspaper advertisement and/or a letter distributed in public elementary schools asking for volunteers to participate in a study of the effects of television on children.[3] The economic and ethnic backgrounds of the subjects were widely varied. Although these characteristics were not considered in assigning subjects to treatment conditions, inspection of the data showed adequate distribution across conditions.

Upon arrival the parent and child were separated, the experiment was explained, and consent was obtained from the parent. Each child was then taken to a room; a television set was turned on, and the child was told he might watch for a few minutes until the experimenter was ready for him. The child watched television alone in the room for about six and one-half minutes. For all groups, the first two minutes consisted of two one-minute commercials videotaped from commercial television during 1970. The commercials (a towel advertisement and an advertisement for a family motion picture film) contained a number of sudden events, presumably intended to be humorous. In one, for example, the father of the household accidentally knocked the dishes off the breakfast table while mother watched with resignation. In the other an unoccupied car accidentally rolled off a pier into the ocean. These commercials had been selected to provide a chance to evaluate the child's facial responsivity to visual inputs. After the commercials, the children in the experimental group were shown the first three and one-half minutes of a program from the series *The Untouchables*. This sequence contained a killing, a chase, the shooting and death of one villain, and an extended fist fight involving the second villain. In contrast, children in the control group viewed a highly active, competitive three-and-one-half-minute sports sequence showing both males and females in races, high jumps, and so on. For all subjects, the final seconds of the program showed a commercial for automobile tires. Within 15 seconds after the commercial began, the experimenter entered the room, told the child it was time to begin, and took the child to the next room.

The child was seated at a response box apparatus modeled after the one employed by Mallick and McCandless (1966). The box displayed a red HURT button, a green HELP button, a white light, and several wires running to a vent in the wall. The subject was told that another child in the adjacent room was playing a game and that the wires were attached to this other child's game. When the other child turned a handle, the white light would be lit, and the subject could press either the HELP or HURT button. The HELP button would make the handle turn more easily, helping the other child win his game, while the HURT button would make the handle hot, hurting the other child and causing him to release the handle. The subject was told that the longer he held down either button, HELP or HURT, the more effect it would have on the other child. When the subject clearly understood the task, the experimenter left the room. Although the subjects were led to believe there was another child in the other room, in fact there was not. The entire procedure was timed to produce 20 trials (white lights).

After the completion of the twentieth trial, the experimenter reentered the room and led the child to another room designated as the "play room." The room contained three attractive nonaggressive toys and a gun, a knife, and two inflated plastic dolls about three feet tall. The child was told he could play for a few minutes, and the experimenter left the room. Aggressive behavior was rated by an observer who watched the child from behind a one-way vision mirror. At the end of the play period, the experimenter reentered the room and asked the child to recall both the television program and the game he had played. (All subjects included in the data analysis correctly recalled the meaning of the red and green buttons and the essential content of the television program they had seen.[4])

A videotape recording of each child's face was made during the entire period the child was in the television room. None of the children included in the data analysis noticed the camera, which was in another room and aimed through a small pane of clear glass. The videorecording also included a multiplexed small image of the television program the child was watching. The sound track of the television program and any sounds made by the child were recorded on the audio track of the videotape.

Three samples from the videotape record of each subject were selected for data analysis. These samples consisted of one from the initial commercials and two from either the violent program or the sports program. The decision about sampling points was *a priori.* The two violent episodes from *The Untouchables* sequence were central to the plot and occupied about the same amount of time (27-30 seconds); as such they were among the longest incidents in the violent program. These two episodes—the shooting and the fight—also depicted two different types of violent actions. In the shooting, a plainclothes policeman fired a pistol from a considerable distance during a chase at a victim who had already been established as a villain by his earlier physical assault on another uniformed policeman. The shooting also showed the consequence of the violence. The victim staggered, fell, and, as the camera zoomed in on his face, said his last few words and died with his pursuers hovering over him. The fight episode involved close body contact among four people, the use of fists, with action so rapid it was hard to be certain who was attacking whom, ending with one person being held physically immobile by three others, including a policeman. The shooting immediately preceded the fight; the fight was the last episode in the first three and one-half minutes of *The Untouchables*.

Two 25-second episodes were selected, from comparable time points in the athletic program, of events which seemed to be of maximal potential interest and suspense. Sports A was an entire women's 100-meter race; Sports B was the final three-quarters of a men's 400-meter relay race. In addition, a thirty-second sample was selected from the initial commercials: about 15 seconds showing the last and most extreme act of carelessness by the husband—knocking all the dishes off the breakfast table, the resignation of his wife, and about fifteen seconds of the other commercial, in which an apparent torpedo fired at a boat is discovered to be a salami.

The measurements of emotion shown in the facial expression during these samples were obtained by requesting eight groups of observers to look at the samples and judge the emotions manifest in each.[5] Each group was composed of about 25 college undergraduates who were paid for their participation. The inserts of the actual television programs (which had been multiplexed onto the tapes when they were recorded) were removed so that the observers could not see the program the child had been watching, but only the child's facial behavior. There were too many samples to show to any one observer—three samples for each of 65 children (Commercial, Shooting and Fighting, or Commercial, Sports A and Sports B). Furthermore, it was preferable that an observer not see more than one sample of any single child, since his judgment of the second or third sample might be influenced by his judgment of the first sample of that child. To solve these problems, four separate videotapes were made, to be shown to four separate groups of observers. Each of these tapes contained about 50 randomly ordered samples; each child appeared only once on each videotape. Each videotape contained about the same number of samples from the commercial, violence, and sports sequences, and each contained about the same number of male and female children. There was a fifteen-second pause between each sample of facial behavior, during which the observer wrote down his or her judgment of emotion. Each of the four videotapes required about fifty minutes for viewing.

There are two schools of thought about ways of measuring judgments of emotion from facial behavior. One approach has emphasized the use of a small set of scales, most commonly including pleasantness and arousal or activity. The other approach has utilized a number of separate emotions, usually distinguishing among various unpleasant feelings (anger, fear, disgust, sadness) and providing a category for positive feelings (happiness). Two different forms were employed in this experiment to permit the use of both of these approaches. One form asked the observers to record their judgments on each of four nine-point scales labeled *pleasant-unpleasant, interested-disinterested, aroused-unaroused,* and *involved-uninvolved.* The other form asked the observers to rate each of the following emotions: *anger, happiness, disgust, fear, pain, sadness,* and *surprise,* using a nine-point scale to record whether each emotion was absent or slightly, moderately, or extremely shown. The observers who used this form were told to consider the duration of a facial expression, its relative intensity, and the frequency of its occurrence in determining whether to judge the emotion as slight, moderate, or extreme. Eight groups of observers were used to obtain independent judgments on each form for the four videotapes.

RESULTS

Facial measures

The pooled judgments of emotion made by the observers were used to establish each child's emotional reactions during each of the three videotape samples. The basic datum was the mean on an emotion scale, calculated across a group of observers who viewed a particular videotape of a child's facial expression and used a scale to record their judgments. For each child there were 33 such means: eleven emotion scale means *(pleasantness, interest, involvement, arousal, anger, disgust, fear, happiness, pain, sadness,* and *surprise*) for each of the three videotape samples (either Commercial plus Shooting and Fighting or Commercial plus Sports A and Sports B). Inspection of the distribution of these means suggested square-root transformations to normalize the distributions. All of the reported results were obtained using transformed scores.[6]

Technically it would be correct to refer to these scores on the children's reactions to the television programs as *judged* emotion, since the data source was observers' judgment of the facial behavior, not the child's own reported emotion nor any direct behavioral measurement. For ease of reading, however, we will forego the qualification *judged,* and more simply use the word *emotion.*

Measures of aggression

Four measures of postviewing aggressive behavior were utilized from the data collected by Liebert and Baron (1971)—three from the button press task and one from the play room situation. The cumulative time that the child held down the red button across all trials seemed to offer a simple measure of the extent to which the child attempted to hurt the other child; this will be referred to as the HURT score. The cumulative time the child held down the green button offered a parallel measure of how much of the time the child attempted to help the other child; this will be referred to as the HELP score. The trial in which the child first pressed the red HURT button was also employed as an indicator of how ready the child was to hurt the other child; this will be referred to as the SLOWNESS TO HURT score. The AGGRESSIVE PLAY measure was the score provided by Liebert and Baron based on time-sampled

Table 1: Intercorrelations among postviewing behavior measures

ACROSS ALL SUBJECTS

	HURT	HELP	SLOWNESS TO HURT
HURT	– –		
HELP	–.21	– –	
SLOWNESS TO HURT	–.60**	.29*	– –
AGGRESSIVE PLAY	–.17	.06	.00

Intercorrelations among postviewing behaviors

	Male aggression group			Female aggression group		
	HURT	HELP	SLOWNESS TO HURT	HURT	HELP	SLOWNESS TO HURT
HURT	– –			– –		
HELP	–.15	– –		–.14	– –	
SLOWNESS TO HURT	–.61*	.42	– –	–.74**	–.01	– –
AGGRESSIVE PLAY	.27	–.27	.02	.22	.04	–.08

Intercorrelations among postviewing behaviors

	Male control group			Female control group		
	HURT	HELP	SLOWNESS TO HURT	HURT	HELP	SLOWNESS TO HURT
HURT	– –			– –		
HELP	.12	– –		–.50*	– –	
SLOWNESS TO HURT	–.60*	.00	– –	–.68*	.58*	– –
AGGRESSIVE PLAY	–.52*	.44	.41	.02	.36	.05

*p < .05
**p < .01

observations of the child during the playroom task.[7] These four scores—HURT, HELP, SLOWNESS TO HURT, and AGGRESSIVE PLAY—were also subjected to the square-root transformation to normalize the distributions.

Table 1 shows the intercorrelations among the four measures of post-viewing behavior. From these relationships it is apparent that although a moderate relationship exists between HURT and SLOWNESS TO HURT, a child's score on any one of the measures is not necessarily informative about his scores on any of the others. Thus all four of the measures were employed in the data analysis.

The lack of a significant correlation between the HELP and HURT measures allowed us to examine whether the same emotional reactions which were positively related to HURT scores might be negatively related to HELP scores. One school of thought views emotional reaction as primarily a change in level of activation. It would be plausible from this viewpoint to predict that emotionally aroused children, regardless of the nature of the arousal (pleasant or unpleasant), would engage in *more* behavior subsequently, regardless of the nature of that behavior (helping or hurting). In contrast, recall that our hypothesis was based upon a different view of emotion, one which emphasizes the specific nature of the emotional arousal as crucial in determining the specific nature of the subsequent behavior. Just as we predicted that children who were *happy, pleasant,* or *interested* when viewing violence would engage in subsequent aggression, we would expect that those who showed *sadness, pain, disgust,* or *fear* would subsequently engage in helping or altruistic behavior. That the HURT and HELP measures were found to be independent supports our view of emotion and allows test of our hypotheses relating specific emotional reactions to different types of subsequent behavior.

Emotion shown while viewing and subsequent behavior: correlations

The simplest way to test our hypothesis and examine the relationship between the two sets of data—scores on emotional reactions shown during the viewing of the television programs and scores on subsequent behavior—was to correlate the two. Significant relationships were predicted between the two sets of scores only among those children who viewed violence, not among those who viewed the sports sequences. Looking *happy* or *sad* while watching a murder was thought to be an index of whether the viewer would subsequently attempt to hurt or help another child, either because the emotional reaction to violence suggested whether the child had identified with victim or aggressor, or because it indicated whether violent activity had been rewarding or punishing to the viewer. But there was no such reason to expect that emotional reactions of *happiness* or *sadness* during a track meet would be correlated with subsequent aggression. Similarly there was no reason to expect that emotional reactions shown during the commercials would relate to subsequent behavior. Correlations were, however, calculated for all groups (those who saw violence and those who saw sports) and for the commercials as well as for the violence or sports sequences, in order to compare the relative occurrence of hypothesized relationships with unexpected relationships. To further assure a conservative evaluation of this investigation as a feasibility study, all statistical outcomes, including those for which directional predictions had been advanced, were evaluated by two-tailed tests.

It has repeatedly been found (Mischel, 1970) that males and females differ in their attitudes about aggression, in their motivations for engaging in aggression, and in the manners or styles in which they manifest aggression. Investigators who have studied the influence of violence viewing on subsequent aggressive behavior have indeed consistently reported differences between male and female subjects. Because of the likelihood of again uncovering such sex-related differences, the investigators decided, before obtaining the data, to conduct the data analysis separately for male and female subjects. We will consider the findings first for the boys, then for the girls.

Table 2 shows all the significant correlations obtained between the 11 emotion scales and the four postviewing behavior measures for the boys. (The complete set of correlations is in Appendix A.) As expected, significant correlations with postviewing behavior were found *only* for the emotions shown by viewers of the violent program. There were no significant correlations between the emotions shown during the commercial and subsequent behavior among either those boys who subsequently saw violence or those who subsequently saw sports. Also, as expected, there were no significant correlations between the emotions shown during the sports sequence and postviewing behavior.

The pattern of correlations shown in Table 2 for the boys strongly supports our hypothesis about the relationship between facial expressions of emotion and subsequent aggressive behavior. There was a lack of perfect correspondence between the findings for the two television violence samples (Shooting and Fighting) and a lack of complete correspondence between the results obtained from the observers who used the four emotion judgment scales (top four rows of the table) and the results from the raters who used the other seven emotion scales (bottom five rows of the table), but over all the consistencies were remarkable.

Positive emotion when viewing violence *(pleasantness* or *happiness)* was positively correlated with the HURT score both when the positive affect was shown during the Shooting and when it was shown during the Fighting. *Sadness* shown during the Shooting was inversely related to the HURT score. *Happiness* shown both during Shooting and during Fighting was positively associated with AGGRESSIVE PLAY. *Interest*

Table 2: Significant correlations between facial measures and postviewing measures—males who viewed violence (N=15)

	Shooting				Fighting			
	HURT	SLOWNESS TO HURT	HELP	AGGRESSIVE PLAY	HURT	SLOWNESS TO HURT	HELP	AGGRESSIVE PLAY
Pleasantness	.60*	--	-.53*	--	.51*	--	--	--
Interest	--	-.72**	-.51*	--	--	-.58*	-.65**	--
Involvement	--	--	--	--	--	-.62*	-.63*	--
Arousal	--	--	--	--	--	-.66**	-.60*	--
Happiness	.67**	--	--	.63*	.55*	--	--	.55*
Anger	--	--	--	--	--	--	--	-.59*
Pain	--	--	.61*	--	--	--	--	--
Sadness	-.61*	--	--	--	--	--	--	--
Surprise	--	--	.60*	--	.56*	-.57*	--	--

*$p < .05$
**$p < .01$

shown during the Shooting was inversely related to the SLOWNESS TO HURT measure (the more *interested* the child, the greater the likelihood that he would be quick to use the HURT button). *Interest, involvement,* and *arousal* shown during fighting were all similarly inversely related to SLOWNESS TO HURT.

The findings on the HELP measure were consistent with the findings for aggressive behavior. *Pleasantness* when viewing the Shooting was inversely related to helping, and *pain* when viewing the Shooting was positively associated with helping. *Interest, involvement,* and *arousal* shown during the Fighting episode were all negatively associated with subsequent helping.

Multiple regression techniques were applied to the data for the boys who viewed violence, in order to reveal which emotions would yield the best prediction of each of the postviewing measures. Generally the results were consistent with those reported in Table 2 and indicate that sets of emotion can account for more than half the variance in predicting postviewing behavior. The details about how these calculations were performed, the results, and the interpretation are given in Appendix B.

The findings on the girls failed to support the hypothesis. Few significant correlations were obtained; more were obtained in response to the commercial than in response to the violent episodes. Table 3 shows the significant correlations for the girls who saw the violent program. It seems wisest not to attempt to interpret the findings on the girls, but instead to view the few correlations which were significant as a chance finding. Only one of the 88 correlations for the reactions to the violence episode was significant. (By comparison, 21 of 88 were significant for the boys.) Further, the commercial produced a few more significant correlations than the program material, for both those who subsequently saw violence and those who saw sports. But even here, the number of obtained significant correlations does not exceed what might occur by

Table 3: Significant correlations between facial measures and postviewing behavior–girls who viewed violence (N=15)

	Videotape sample		
	Commercial		Shooting
	HURT	SLOWNESS TO HURT	HURT
Pleasantness	–.53*	– –	– –
Happiness	–.47*	– –	– –
Pain	– –	– –	.47*
Sadness	.62**	–.63**	– –

*$p < .05$
**$p < .01$

chance, and, unlike the data for the boys, they do not follow the predicted pattern.

Attempting to understand why the emotion shown during the viewing of violence correlated with postviewing behavior among the boys but not among the girls, we considered two possible explanations but largely dismissed them after further data analysis.

One possibility was that perhaps the girls did not experience the same emotions as the boys when viewing the violent program. If, for example, the girls showed little or no *pleasantness, happiness, sadness*, or *interest* (emotions which correlated with subsequent behavior for the boys), then their emotions could not correlate with their postviewing behavior, since correlations require that there be a range of scores on each variable being related. Analysis of variance (ANOVA) techniques tested this possibility by determining whether there was any difference between the emotion shown by the girls and that shown by the boys, regardless of the program viewed or particular program episodes. Three of the 22 ANOVAs revealed a sex difference; on the Fighting sequence and Sports B (the men's relay race), boys were more *involved* and *aroused* than girls, and on *sadness* there was an interaction between sex and type of program (but no main effect for sex). However, these differences between boys and girls are not sufficient to explain why emotional reactions were correlated with postviewing behavior for boys but not for girls. While boys and girls did differ in *involvement, arousal*, and *sadness* in response to the Fighting episode, they did not differ in *interest, pleasantness, happiness, anger*, and *surprise*; yet all these latter emotions did correlate with postviewing behavior for boys and not for girls. Furthermore, boys and girls did not differ in *any* of their emotional reactions to the Shooting episode, yet many of these emotions were correlated with subsequent behavior for boys but not for girls. In general, then, the suggestion that girls did not experience the same range of emotion as the boys—and therefore that their emotional reactions could not be similarly correlated with their subsequent behavior—does not appear tenable. (Appendix C reports these analyses in detail.)

A second possible explanation was that perhaps the girls showed much less aggressive behavior than the boys. If that were so, then the emotions shown by the girls could not be correlated with subsequent aggressive behavior because correlations require (as noted earlier) a range of scores on both variables. ANOVA failed to find an overall sex difference or a sex difference in relation to the type of program viewed (sports or violence) for the HURT measure, the HELP measure, or the SLOWNESS TO HURT measure. The ANOVA for the AGGRESSIVE PLAY measure did reveal a significant effect for sex and for program viewed in relation to sex; the boys who had watched a violent program showed more AGGRESSIVE PLAY than the girls who had watched violence, while there was no such difference on this measure between boys and girls who had watched the sports sequence. (Appendix D reports these analyses.) Despite this difference in AGGRESSIVE PLAY, the fact that the boys did not differ from the girls in the three measures of button press task behavior requires that we reject the possibility that a sex difference in postviewing behavior can account for the fact that emotional reactions during the viewing of television correlated with postviewing behavior for boys and not for girls.

Summary of sex differences: Two possible differences between boys and girls (emotional reactions and postviewing behavior) were examined as possible bases for the finding that emotional reactions to the programs were correlated with postviewing behavior for boys but not for girls. The data analysis failed to find substantial differences between boys and girls on either. Two other possibilities, which could not be tested in this experiment, will be considered in the discussion.

Let us turn now to another method of examining the data relevant to our hypothesis about emotional reactions and post viewing aggression.

Emotion while viewing, program viewed, and sex of the viewer as predictors of postviewing behavior: analysis of variance

Analysis of variance techniques provided a complementary pattern of results as well as new information which had not been revealed in the correlational analysis reported in Tables 2 and 3. The ANOVAs showed that children who had different emotional reactions to the violent program showed different behavior on some the postviewing measures. (The correlations could determine only that emotional reactions were related to postviewing measures, not whether children who showed different emotional reactions to the violent program subsequently manifested significantly different aggressive behavior.) The ANOVAs also revealed that when children who had shown a similar emotional reaction were compared, those who had seen the violent program subsequently engaged in more aggressive behavior than those who had seen the sports program.

Three independent variables were used in calculating the ANOVAs: program viewed (violence or sports), sex of the viewer, and emotion shown (a high, medium, or low score based upon trichotomizing the scores on an emotion scale within each group of viewers). ANOVAs were calculated utilizing each of the emotions and each of the postviewing behavior measures for which a significant correlation had been obtained among the boys who viewed violence (Table 2). Appendix E reports the results of these separate ANOVAs. A summary of these findings, integrating the results across separate ANOVAs, is presented below.

Differences among those who viewed violence. Among the boys, those who reacted to the violence with the most *pleasantness* (high scores) subsequently engaged in more hurting behavior than those who reacted to the violence with the most *unpleasantness* (low scores). Similarly, those who reacted to the violence with the least *sadness* subsequently engaged in more hurting behavior than those who reacted with the most *sadness.* These results were found for both the Shooting and the Fighting samples. Further, among the boys, those who reacted to the Shooting with the most *disinterest* subsequently were slower to hurt and engaged in more helping behavior than those who showed the most *interest.*

No such findings were obtained for the girls, and no differences were found for either boys or girls in AGGRESSIVE PLAY as a function of emotional reactions.

Differences among those who viewed violence or sports. Among the boys who showed the most *pleasant* reactions, those who viewed violence subsequently engaged in more hurting behavior than those who viewed sports. Similarly, among the boys who showed the least *sadness,* those who viewed violence subsequently engaged in more hurting behavior than those who viewed sports. These findings were obtained both when reactions to the Shooting were compared with reactions to Sports A and when reactions to the Fighting were compared with reactions to Sports B. Among the boys who showed the most *disinterest,* those who viewed Sports A were slower to hurt than those who viewed the Shooting.

Among the girls who showed the most *unpleasantness,* those who viewed violence subsequently engaged in more hurting behavior than those who viewed sports. Similarly, among those girls who showed the most *sadness,* those who viewed violence subsequently engaged in more hurting behavior than those who viewed sports. These findings were obtained both when reactions to the Shooting were compared with reactions to Sports A and when reactions to the Fighting were compared with reactions to Sports B.

No significant findings were obtained as a function of emotional reaction, on either the HELP or AGGRESSIVE PLAY measures, for either boys or girls.

Differences between boys and girls. Among the children who showed the most *interest* when viewing the Shooting, the girls were slower to hurt than the boys.

DISCUSSION

This experiment was designed as a feasibility study, to determine if facial expressions of emotion shown during the viewing of televised violence predict subsequent behavior. The success achieved with the boys suggests a positive answer to the feasibility question. The results obtained with the five- and six-year-old boys confirmed the specific hypothesis: boys whose facial expressions during the Shooting or the Fighting episodes showed *happiness, pleasantness,* and not *sadness,* tended to use the HURT button more than boys whose facial expressions showed *unpleasantness, sadness,* and not *happiness.* Boys whose facial expressions during the Shooting or Fighting showed *interest* used the HURT button earlier in the task than boys whose facial expressions during the Shooting or Fighting showed *disinterest.*

The facial expressions shown during violence were also related to the use of the HELP button. Boys whose facial expressions during the Shooting showed *unpleasantness, pain, surprise,* or *disinterest* tended to use the HELP button more than did those whose facial expressions showed *pleasantness,* little *pain,* no *surprise,* or *interest.*

The facial expressions shown by the boys during the violent program were also related to aggression shown in the other task, the play situation. Boys whose facial expressions during the Shooting or Fighting episodes showed *happiness* tended to engage in more aggressive play than those boys whose facial expressions during the Shooting and Fighting were not *happy.*

The findings for the five- and six-year-old boys also showed that it was not simply the emotional reaction to television which predicted subsequent aggression, but more specifically the emotional reaction to a violent program. Emotional reactions to the sports program or to the commercials did not relate to subsequent aggression. In addition, there were no differences in the emotional reactions shown to violence as compared to the sports program on most of the emotion scales correlated with postviewing behavior. Another way to illustrate this part of the findings is to examine just those boys who showed the most *pleasantness* and the least *sadness.* Among these boys, those who were viewing violence behaved more aggressively afterwards than those who were viewing the sports event.

In presenting the results on the correlations among the postviewing behavior measures, we noted that the lack of relationship between the HURT and HELP measure was important since it would allow a test of an alternative conceptualization of the relationship between emotion and postviewing behavior to the one we had hypothesized. That alternative view might lead to the hypothesis that when children are emotionally aroused (regardless of whether the arousal is positive or negative), they will subsequently engage in more active behavior of any kind, hurting *or* helping. The findings from this experiment directly contradict this hypothesis. The *arousal* measure was *not* positively related to either HURT or HELP measures; instead *arousal* during the Fighting episode was inversely related to the SLOWNESS TO HURT measure and to the HELP measure, and unrelated to the HURT measure. The direction or type of emotional arousal predicted the direction or type of subsequent activity. That is, those boys who appeared *pleasant, happy*, and not *sad* showed more hurting and less helping behavior; those who appeared *unpleasant*, not *happy*, and *sad* showed more helping and less hurting behavior. As we hypothesized, it was the *type* of emotional reaction which was crucial.

These findings for the boys suggest that, in order to understand the influence of televised violence on the likelihood of subsequent aggression, we must consider the child's emotional reaction to the violence. While across the population seeing violence may increase the likelihood of aggression, an equally important predictor may be the viewer's reaction to violence. This study has shown that there are markedly different emotional reactions to the same violent program and that these different reactions, as indicated by facial expression, *do* predict subsequent aggression. Put another way, the results on the boys suggest the necessity of considering not just whether the input is violent or not, but also the perceiver's reactions to that input, if the variations in the amount of subsequent aggression among those who view violence are to be understood.

It is, of course, necessary to regard the implications of this study as suggestive rather than conclusive. The experiment should be replicated with children of the same and other ages, with other violence inputs, and perhaps also in other settings.

Let us now discuss the failure to obtain results with the girls. We considered two possible explanations of the fact that emotional reactions shown during the viewing of violence were correlated with subsequent aggression for the boys but not for the girls. From the findings it appears unlikely that this is due to a difference between boys and girls either in emotional reactions during a violent program or in subsequent aggression. The data do not allow tests of other possible explanations, but at least two which could be explored in further research can be mentioned. First, it may have been important that the violence shown was male, not female—male actors, both as victims and as aggressors; male roles, cops and robbers; male forms of aggression, shooting and fist fighting. Second, it may be necessary to consider individual differences in the extent to which girls have been censored or punished for empathizing with or approving of aggressive activity. Other studies are necessary to test these possibilities, utilizing other types of inputs with female characters and with other types of violence, using girls of different ages and differentiated in terms of attitudes about aggression.

The findings on the boys raise the tantalizing question of why some five-year-old boys react with *happiness* and *interest* while others react with *sadness* and *disinterest* to a shooting or to fighting. The question is important, since these different emotional reactions shown in facial expressions did serve to account in some large part for differences in subsequent aggression. The experiment was not designed, however, to examine the determinants of individual differences in emotional reaction to the viewing of violence. All that can be done is to mention some of the possibilities which should be explored in subsequent research. The difference between a positive and a negative emotional reaction to the viewing of televised violence may be due to transient phenomena such as the mood or expectations of the viewer, the type of program watched on television prior to the experiment, and so on. Or the difference may be related to such enduring characteristics as personality, parental attitudes toward aggression, and child rearing practices. Further research is needed to explore these possibilities. Further study should also study the stability of the child's emotional reaction to the viewing of televised violence: is it consistent over repeated occasions, different program formats, different forms of violence?

In the design of this feasibility study, one aim was to compare two different methods for extracting information about emotion from facial expression. Only one of these methods, the grosser and less costly, has been used—aggregated judgments by observers of the facial expressions. The success with this gross measure suggests that applying the more precise measurement system would be fruitful. The Facial Affect Scoring Technique (Ekman, Friesen, and Tomkins, 1971) would answer new kinds of questions. Exactly when is a particular emotion shown; does a child show a *happy* facial expression when the aggressor shoots, when the victim falls, as the victim dies, after the violence subsides? What is the sequence of emotional reactions; do some children show *happiness* at the attack, followed by *sadness* at the death, followed by *happiness* as the violence subsides? This more precise and exact measurement procedure might even succeed in predicting the subsequent aggression of the girls. If girls (more than boys) could be expected to have conflicting attitudes about violence, then a measurement procedure which looks at sequences of emotional reactions could be useful. But this is another question which can be answered only through more study.

The implications of this experiment for both basic and applied research questions should be noted. For students of emotion, and in particular for those interested in facial expression of emotion, this experiment provides the first evidence that the momentary facial expression shown during an emotional experience can predict subsequent complex social behavior. In the long history of research on whether facial expressions provide accurate information about others, research has been *post*dictive, attempting to determine whether the eliciting stimulus could be inferred from the facial expression. This experiment has fundamental importance in showing the *pre*dictive value of facial expressions of emotion.

The experiment also offers promise for a number of applied studies within the area of television violence and aggression. This experiment has shown that children differ in their emotional reactions to violence and that these differences in emotional reaction do predict, for the boys, differences in subsequent aggressive behavior. If further research confirms the finding that measures of facial expression while viewing violence predicts subsequent aggressive behavior, and determines the generality of that finding with different age groups, with girls as well as boys, and with different types of violent material, then this measure should be included in further research on the host of questions relevant to understanding television and social behavior. The interactions among input, setting, prior exposure, and viewer variables in relation to subsequent aggression may be better illuminated if emotional reactions as revealed by facial expressions are also considered. The particular utility of measuring facial expressions is that records can be gathered unobtrusively, without influencing the subject, and from very young subjects who may not be accessible to verbal testing or for whom it may be difficult to arrange a postviewing aggressive task situation.

In final summary, this study points to several important conclusions:

(1) It does appear feasible to obtain unobtrusively valuable indicators of emotional state from the facial expressions of children watching television.

(2) These facially revealed emotional states appear to be important predictors of later significant social behavior.

(3) In all, several different emotions seem relevant in predicting later behavior. *Pleasantness* and/or *happiness, interest, involvement, arousal, anger, pain, sadness,* and *surprise* were each significantly correlated with at least one class of postviewing behavior.

(4) Similarly, quite varied and relatively uncorrelated social behaviors can be predicted from emotional states. While altruistic (HELP) behavior cannot be predicted from aggressive (HURT) behavior (the intercorrelation being only -.21), both kinds of behavior can be predicted above chance from the emotional states revealed during the preceding viewing of televised violence (at least for the boys).

(5) Emotional states while watching violence account for a significant share of the variance in predicting later aggression and altruism. For boys, multiple correlations among a few emotions can account for more than 50 percent, and up to 75 percent, of the variance on all four postviewing behaviors.

(6) While viewing televised violence may increase the likelihood of aggressive behavior, an equally important factor may be the emotional response of the child to dramatic portrayals of aggression. Thus, the interaction of input and affective response in understanding the impact of television on children appears to be a uniquely promising and important prospect for future research.

FOOTNOTES

1. We are grateful to Norma McCoy Irons, Associate Professor of Psychology, San Francisco State College, for her many valuable contributions to the design of this study and interpretation of results. The research was conducted under a contract with the National Institute

of Mental Health. Drs. Ekman, Friesen, Zlatchin, and Malmstrom are at the Department of Psychiatry (Langley Porter Institute), University of California, San Francisco. Dr. Liebert is at the Department of Psychology, State University of New York at Stony Brook. Dr. Harrison is at the Department of Communication, Michigan State University. Dr. Baron is at the Department of Psychology, University of South Carolina.

2. Videorecordings of facial expressions were obtained also on a comparable number of eight- and nine-year-old children, but these records have not been analyzed.
3. A more detailed description of the subject population and the experimental procedures is given in Liebert and Baron, 1971.
4. Nine children were terminated before the data were collected because they refused to remain alone, cried, or left the experimental situation. Five other children did not understand or follow instructions for the button press task and three played or explored the room instead of watching television. The remaining subjects totaled 30 boys and 35 girls, as reported in the text.
5. The validity and reliability of data gathered from untrained observers is strongly indicated in recent research (Ekman and Friesen, 1971; Ekman, Friesen, and Tomkins, 1971; Ekman, Friesen, and Ellsworth, 1971; Ekman, in press): (a) evidence indicates that the average young adult of college age and intelligence can effectively discriminate facial expression of emotion, although he may not be able to specify the facial cues or the inference process he employs; (b) individual differences in ability to perform the task or to judge a specific affect can be neutralized by pooling the judgments of a group of observers, to yield a reasonable measure.

   In the present study, with two groups of observers judging the same tapes, it was possible to obtain some indication of relative reliability by comparing judgments on two similar scales, *pleasantness-unpleasantness* and *happiness*. Judgments on these scales are only partially redundant. While the opposite end of the pleasantness scale was *unpleasantness*, the opposite end of the happiness scale was *not showing happiness*, rather than specification of some unhappy emotion (anger, sadness, etc.). Nevertheless, the correlations ranged from .65 to .88, with half the correlations above .80. These findings show that at least in terms of happiness and pleasantness, judgments were reliable, in that similar groups of observers provided similar measurements of the same stimuli.
6. This statistical adjustment, which is common practice for highly skewed data, was selected on *a priori* grounds according to the recommendations of Winer (1962, pp. 218-22). Specifically, the transformation $x' = \sqrt{x} + \sqrt{x+1}$ was employed.
7. Of the other measures which could be derived from the button press task we used three: number of HURT button presses, mean HURT

duration and mean HELP duration. The pattern of findings was generally the same as is reported.

REFERENCES

Ekman, P. Universals and cultural differences in facial expressions of emotion. In Cole, J. K. (Ed.) *Nebraska symposium on motivation, 1971.* Lincoln, Neb.: University of Nebraska Press (in press).

Ekman, P., and Friesen, W. V. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 1971 **17**(2), 124-29.

Ekman, P., Friesen, W. V., and Ellsworth, P. *Emotion in the human face: guidelines for research and a review of findings.* New York: Pergamon Press, 1971.

Ekman, P., Friesen, W. V., and Tomkins, S. S. Facial affect scoring technique: a first validity study. *Semiotica*, 1971, **3**(1), 537-58.

Ekman, P., Sorenson, E. R., and Friesen, W. V. Pan-cultural elements in facial displays of emotions. *Science*, 1969, **164**(3875), 86-88.

Liebert, R. M., and Baron, R. A. Short-term effects of televised aggression on children's aggressive behavior. In *Television and social behavior*, Vol. 2 (this series). Washington, D. C.: U. S. Government Printing Office, 1971.

Mallick, S. K., and McCandless, B. R. A study of catharsis of aggression. *Journal of Personality and Social Psychology*, 1966, **4**, 591-96.

Mischel, W. Sex typing and socialization. In Mussen, P. (Ed.) *Carmichael's manual of child psychology*, Vol. 2 (3rd Ed.). New York: Wiley, 1970.

Tomkins, S. S. *Affect, imagery, consciousness*, Vol. 1: *The positive affects.* New York: Springer, 1962.

Tomkins, S. S. *Affect, imagery, consciousness*, Vol. 2: *The negative affects.* New York: Springer, 1963.

Winer, J. B. *Statistical principles in experimental design.* New York: McGraw-Hill, 1962.

Appendix A: Correlations between facial measures and postviewing measures for males who viewed violence

Table A-1 presents the correlations between postviewing behaviors and those affects which showed at least one significant relationship with some postviewing measure. Table 2, in the text, displayed this data in an abbreviated form, highlighting the significant correlations only.

As noted in the text, the correlational hypotheses were tested using two-tailed tests, although the investigators did have specific, directional hypotheses for some of the emotions, e.g., the prediction that *happiness* would correlate with later aggression. If the more generous, one-tailed significance level is used, four additional relationships become "significant": (a) the greater the involvement while watching the shooting sequence, the quicker the individual is to aggress; (b) the more pain expressed during the shooting sequence the slower the individual is to aggress; (c) the more sadness shown during the shooting, the less aggressive play; and (d) the more pleasantness shown during the fighting sequence, the less the individual helps later. These findings conform to expectations and are consistent with the other findings discussed in the text.

Table A-1: Correlations between facial measures and postviewing measures—males who viewed violence (N=15)

	SHOOTING				FIGHTING			
	HURT	SLOWNESS TO HURT	HELP	AGGRESSIVE PLAY	HURT	SLOWNESS TO HURT	HELP	AGGRESSIVE PLAY
PLEASANTNESS	.60*	–.36	–.53*	.23	.51*	–.24	–.44Δ	.39
INTEREST	.28	–.72**	–.51*	–.01	.34	–.58*	–.65**	–.13
INVOLVEMENT	.09	–.49Δ	–.30	.18	.32	–.62*	–.63*	–.14
AROUSAL	.33	–.14	.31	.20	.31	–.66**	–.60*	–.14
HAPPINESS	.67**	–.31	–.31	.63*	.55*	–.10	–.19	.55*
ANGER	–.43	.24	.13	–.43	–.01	–.15	.10	–.59*
PAIN	–.19	.45Δ	.61*	–.08	.10	–.35	–.07	–.14
SADNESS	–.61*	.28	.31	–.47Δ	–.16	–.24	–.04	–.21
SURPRISE	.30	–.09	.60*	.13	.56*	–.57*	–.14	.35

*$p < .05$ (two-tailed test)
**$p < .01$ (two-tailed test)
Δ$p < .05$ (one-tailed test)

Appendix B: Multiple regression analysis

Multiple correlations were calculated only for the data on the boys who had seen the violent program, to reveal which emotions would yield the best prediction of each of the postviewing measures and to indicate how large the correlation would be when more than one emotional reaction to a television episode was considered. In calculating the multiple regressions, the judgments from each group of observers were first considered separately and then again pooled. Table B reports only those multiple correlations in which the figure obtained was larger than the figures reported in Table 2 when each of the emotion scales were separately correlated with the postviewing measures. Further, following the usual statistical conventions no more than the first three scales were reported, and only if there was an increase in the correlation by at least five points between the first and the second, or the first and the third, scale. Finally, the correlations reported are corrected for the small sample size; this correction acts to reduce the size of the coefficient.

Sizeable multiple correlations were obtained. *Pleasantness* and *arousal* during the shooting predicted the HURT scores. Not being *aroused*, not being *surprised* and being *happy* predicted SLOWNESS TO HURT. Not being *angry*, being *happy* and being *sad* predicted AGGRESSIVE PLAY; the contribution of *sadness* in this multiple correlation was the only finding which contradicted the findings from Table 2, and it is not readily explicable. Showing *pain* and *surprise* but not *happiness* during the shooting correlated with the HELP measure. Not being *interested* but showing *disgust* during the fight correlated with the HELP measure.

Table B-1: Multiple correlations of emotions during viewing and postviewing behavior—males who viewed violence (N=15)

TV episode	Emotions	R	Postviewing measure
Shooting	PLEASANTNESS	.56*	HURT
	PLEASANTNESS + AROUSAL	.77**	
Fighting	–AROUSAL	.62*	SLOWNESS TO HURT
	–AROUSAL - SURPRISE	.66**	
	–AROUSAL - SURPRISE + HAPPINESS	.71*	
Fighting	–ANGER	.54*	AGGRESSIVE PLAY
	–ANGER + HAPPINESS	.77**	
	–ANGER + HAPPINESS + SADNESS	.87**	
Shooting	PAIN	.57*	HELP
	PAIN + SURPRISE	.68*	
	PAIN + SURPRISE –HAPPINESS	.80**	
Fighting	–INTEREST	.62*	HELP
	–INTEREST + DISGUST	.71*	

*p = .05
**p = .01

Appendix C: An examination of differences in emotional reactions between boys and girls

These analyses were performed in order to determine whether there were differences in the emotional reactions to the violent or sports programs between boys and girls which might account for the finding (Table 2) that emotional reactions correlated with postviewing measures for boys but not for girls. Eleven separate ANOVAs were calculated, one for each of the emotion scales on the behavior shown during the viewing of the Shooting and Sports A; and another eleven ANOVAs were calculated for the emotion scores on the behavior shown during the viewing of the Fighting and Sports B.

Only one of the eleven ANOVAs on the Shooting and Sports A facial behavior scores yielded a significant finding. Table C-1 shows that when the *happiness* scores were considered, there was more *happiness* shown during the sports than violent sequence regardless of the sex of the viewer ($p = .05$). Three of the eleven ANOVAs on the Fight and Sports B sequence, however, yielded a significant finding. Table C-2 shows that when *sadness* scores were employed, there was a difference between the reactions of the boys and girls to the violence and sports sequences (sex of the viewer x program viewed interaction $p = .05$). Males showed more *sadness* during the Fight than females (t-test, $p = .01$), while there was no difference between males and females in the *sadness* shown during the Sports B sequence. Tables C-3 and C-4 show that during Fighting and Sports B (the men's relay race), there were main effects for sex on *involvement* and *arousal*—i.e., boys were more aroused and involved than girls while watching both the fight scene and the 400-meter relay race.

Since only four of the 22 ANOVAs show any significant differences—and only three of these indicate any sex difference—these findings must be interpreted with caution. In general, however, the results appear to suggest a negative answer to the substantive question: did the girls and boys differ in their emotional reactions during violent and sports programs? When emotions shown while watching the Shooting and Sports A sequence were considered, there were *no* significant differences between boys and girls on any of the eleven emotions, regardless of the program viewed. When emotions shown during the Fighting and Sports B sequence were considered, there were no differences between boys and girls on *pleasantness, interest, happiness, anger, disgust, pain, surprise,* or *fear*, but the boys showed more *sadness* than the girls during the Fight and greater *involvement* and *arousal* during both the Fight and Sports B.

The possibility of a difference between males and females in the emotions shown during the television programs was raised as one conceivable explanation of why significant correlations were obtained between emotions and postviewing behavior for boys but not girls. These ANOVAs suggest that by and large there were few differences in the emotions shown by boys and girls; therefore this explanation seems doubtful.

Table C-1: The differences on HAPPINESS for boys and girls watching "Shooting" or "Sports A"

Source	df	MS	F
Shooting vs Sports A (A)	1	2.07	4.23*
Boys vs Girls (B)	1	0.12	
Interaction: A x B	1	0.01	
Error	61	0.49	

Table C-2: The differences on SADNESS for boys and girls watching "Fighting" or "Sports B"

Source	df	MS	F
Fighting vs Sports B (B)	1	0.01	
Boys vs Girls (B)	1	0.13	
Interaction: A x B	1	1.04	4.23*
Error	61	0.25	

*p < .05

Table C-3: The differences on INVOLVEMENT for boys and girls watching "Fighting" or "Sports B"

Source	df	MS	F
Fighting vs Sports B (A)	1	0.14	
Boys vs Girls (B)	1	2.70	4.85*
Interaction: A x B	1	0.01	
Error	61	0.56	

Table C-4: The differences on AROUSAL for boys and girls watching "Fighting" or "Sports B"

Source	df	MS	F
Fighting vs Sports B (A)	1	0.00	
Boys vs Girls (B)	1	2.23	4.55*
Interaction: A x B	1	0.08	
Error	61	0.49	

*p < .05

Appendix D: An examination of differences in postviewing behavior between boys and girls

These analyses were performed in order to determine whether there were differences in the postviewing behavior of boys and girls which might account for the finding (Tables 2 and 3) that emotional reactions correlated with postviewing behavior for boys and not for girls. ANOVAs were calculated utilizing each of the emotions and each of the postviewing behavior measures for which a significant correlation had been obtained for the boys who viewed violence (Table 2). Each of these ANOVAs was a 2 x 2 x 3 analysis: sex (male vs. female), program (sports vs. violence), emotion (high, medium, and low). Appendix E reports the findings dealing with the emotion variable; here we will only consider the findings on the sex variable, or interactions of sex with program viewed. A total of 19 ANOVAs were performed, six utilizing the HURT scores, four with the SLOWNESS TO HURT scores, six with the HELP scores, and three with the AGGRESSIVE PLAY scores.

There were no significant findings for the sex of the viewer, or for sex of the viewer in relation to the type of program viewed, for the ANOVAs on the HURT, HELP, or SLOWNESS TO HURT measures. The ANOVAs for the AGGRESSIVE PLAY measures, however, showed a significant effect for sex and for program viewed in relation to sex. Tables D-1–D-3 show the ANOVAs for the AGGRESSIVE PLAY measures when conducted with the emotion scales of *happiness*—on both Shooting vs. Sports A and Fighting vs. Sports B—and *anger* on Fighting vs. Sports B. In all three, the boys showed more AGGRESSIVE PLAY than the girls.

The possibility of a difference between boys and girls in postviewing behaviors was examined as an explanation of why the emotions shown during the violent program were correlated with postviewing behavior for boys but not for girls. The fact that the ANOVAs failed to reveal any sex difference for the three button press task measures (HURT, SLOWNESS TO HURT, and HELP), even though behavior on these measures was correlated with emotions shown for the boys, suggests that an explanation which postulates a difference in postviewing behavior between boys and girls must be rejected.

Table D-1: The emotion HAPPINESS displayed by boys and girls
watching "Shooting" or "Sports A",
related to AGGRESSIVE PLAY

Source	df	MS	F
Shooting vs Sports A (A)	1	15.63	8.08**
Boys vs Girls (B)	1	19.77	10.23**
High, medium or low happiness (C)	2	0.58	
Interaction: A x B	1	10.01	5.18*
Interaction: A x C	2	0.45	
Interaction: B x C	2	1.35	
Three-way interaction: A x B x C	2	2.87	
Error	53	1.93	

*p < .05
**p < .01

Table D-2: The emotion HAPPINESS displayed by boys and girls
watching "Fighting" or "Sports B",
related to AGGRESSIVE PLAY

Source	df	MS	F
Fighting vs Sports B (A)	1	14.43	8.22**
Boys vs Girls (B)	1	18.42	10.49**
High, medium or low happiness (C)	2	1.70	
Interaction: A x B	1	11.01	6.27*
Interaction: A x C	2	4.66	
Interaction: B x C	2	3.55	
Three-way interaction: A x B x C	2	0.34	
Error	53	1.76	

*p < .05
**p < .01

Table D-3: The emotion ANGER displayed by boys and girls
watching "Fighting" or "Sports B",
related to AGGRESSIVE PLAY

Source	df	MS	F
Fighting vs Sports B (A)	1	15.41	8.51**
Boys vs Girls (B)	1	19.53	10.78**
High, medium or low anger (C)	2	0.90	
Interaction: A x B	1	10.18	5.62*
Interaction: A x C	2	3.23	
Interaction: B x C	2	1.02	
Three-way interaction: A x B x C	2	3.80	
Error	53	1.81	

*p < .05
**p < .01

Appendix E: Emotion while viewing, program viewed, and sex of the viewer as predictors of postviewing behavior

In these ANOVAs the measures of postviewing behavior were cast as the dependent variables. There were three independent variables in the 2 x 2 x 3 ANOVAs: sex of the viewer, program viewed (violence vs. sports) and emotion shown, a high, medium or low score based upon trichotomizing the scores on each of the emotions and each of the postviewing behavior measures for which a significant correlation had been obtained for the boys who viewed violence (Table 2).

Table 2 reported three emotions, *pleasantness, happiness*, and *sadness*, correlated with the HURT measure. Six ANOVAs were performed, one for each of these emotion scales, first utilizing the judgments from the videotape samples when the children had been viewing either the Shooting or Sports A, and again for the videotape samples of their facial behavior while they watched either the Fighting or Sports B. Only the former yielded significant findings for the emotion variable.

Table E-1 shows the ANOVA results for *pleasantness*; Table E-2 shows the ANOVA for *sadness*. The ANOVA for *happiness* and the HURT measure failed to yield significant findings for emotion or interactions with emotion, although the tendencies were consistent with the ANOVAs for *pleasantness* and *sadness*. When *pleasantness* was employed in the ANOVA, there was an interaction between the program and the emotion, when *sadness* was employed in the ANOVA, there was an interaction between sex and emotion, and also among emotion, sex and program. Figure 1 shows these findings within these ANOVA classifications. T-tests between means showed the following significant differences:

a. Among males who viewed violence, those high on *pleasantness* had greater HURT scores than those low on *pleasantness* (p = .05); among males who viewed violence, those low on *sadness* had greater HURT scores than males high on *sadness* (p = .01).

b. Among males who were high on *pleasantness*, those who viewed violence had greater HURT scores than those who viewed sports (p = .05); among males who were low on *sadness*, those who viewed violence had greater HURT scores than those who viewed sports (p = .05).

c. Among females who were low on *pleasantness*, those who viewed violence had greater HURT scores than those who viewed sports (p = .05); among females who were high on *sadness*, those who viewed violence had greater HURT scores than those who viewed sports (p = .05).

These findings on the HURT measure were internally consistent and consistent with the correlational data (Table 2).

Table 2 reported that *interest* was correlated with SLOWNESS TO HURT for the males who viewed violence, when both the Shooting sample and the Fighting sample were analyzed. *Involvement, arousal,* and *surprise* were correlated with SLOWNESS TO HURT but only for the Fighting sample. Five ANOVAs were performed utilizing SLOWNESS TO HURT scores as the dependent variable: two for *interest*, in the Shooting and Sports A, and in the Fighting and Sports B; one for *arousal*, one for *involvement*, and one for *surprise*, utilizing only the scores from the Fighting and Sports B samples. Only one of these five ANOVAs yielded significant findings for the emotion variable—*interest* in the Shooting and Sports A sample. Table E-3 shows the results of this ANOVA. When *interest* in the Shooting or Sports A sample was analyzed, there was a main effect for the emotion, and an interaction between emotion, sex and program. Figure 2 shows these findings.

Within these ANOVA classifications, t-tests between means showed the following significant differences:

a. Those males *disinterested* (low) in the Shooting were slower to hurt than those high on *interest* ($p < .05$).

b. Among males who were high on *interest*, those who viewed Sports A were slower to hurt than those who viewed the Shooting ($p < .01$).

c. Among children who were most *interested* (high) in the Shooting, the girls were slower to hurt than the boys ($p < .05$). These findings on SLOWNESS TO HURT were internally consistent and also consistent with the correlational results reported earlier in Table 2.

Table 2 reported that six emotions correlated with the HELP measure: *pleasantness, interest, pain,* and *surprise* during the Shooting and *interest, involvement,* and *arousal* during the Fighting. Only one of the six ANOVAs performed with the HELP measure yielded a significant finding for emotion variable—there was a main effect for *interest* shown in Shooting and Sports A. Table E-4 and Figure 3 show the findings from this ANOVA. The main effect for *interest* means that those more interested in the program, regardless of the content of the program, had smaller HELP scores than those low in *interest*. The figure shows that this appears to be so only for the males who saw violence and the females who saw sports; t-tests showed that the only significant difference among the groups was for the males who saw violence. Those boys highly *interested* in the Shooting had smaller HELP scores than those who showed low *interest* in the Shooting ($p < .05$). When this analysis was performed utilizing the mean HELP score instead of the total HELP score, significant results for emotion were not obtained. This is the only instance in which the results obtained with HURT or HELP scores were not duplicated when mean cumulative scores were employed.

Table 2 reported three emotion scales correlated with AGGRESSIVE PLAY, but none of the ANOVAs utilizing these scales yielded a significant finding for emotion, or emotion in interaction with sex and/or program viewed.

Table E-1: The emotion PLEASANTNESS displayed by boys and girls watching "Shooting" or "Sports A," related to later HURT behavior

Source	df	MS	F
Shooting vs. Sports A (A)	1	124.81	
Boys vs Girls (B)	1	7.02	
High, medium, or low pleasantness (C)	2	89.69	
Interaction: A x B	1	0.00	
Interaction: A x C	2	252.40	4.02*
Interaction: B x C	2	142.00	
Three-way interaction: A x B x C	2	96.46	
Error	53	62.77	

* p < .05

Table E-2: The emotion SADNESS displayed by boys and girls watching "Shooting" or "Sports A", related to later HURT behavior

Source	df	MS	F
Shooting vs Sports A (A)	1	120.87	
Boys vs Girls (B)	1	6.11	
High, medium, or low sadness (C)	2	22.26	
Interaction: A x B	1	0.05	
Interaction: A x C	2	148.56	
Interaction: B x C	2	245.03	4.00*
Three-way interaction: A x B x C	2	194.80	3.18*
Error	53	61.33	

*p < .05

Table E-3: The emotion INTEREST displayed by boys and girls watching "Shooting" or "Sports A", related to later SLOWNESS TO HURT behavior

Source	df	MS	F
Shooting vs Sports A (A)	1	21.39	
Boys vs Girls (B)	1	4.28	
High, medium or low interest (C)	2	29.97	3.69*
Interaction: A x B	1	0.36	
Interaction: A x C	2	6.08	
Interaction: B x C	2	0.89	
Three-way interaction: A x B x C	2	42.44	5.23**
Error	53	8.12	

*p < .05
**p < .01

Table E-4: The emotion INTEREST displayed by boys and girls watching "Shooting" or "Sports A", related to later HELP behavior

Source	df	MS	F
Shooting vs Sports A (A)	1	45.46	
Boys vs Girls (B)	1	52.03	
High, medium or low interest (C)	2	233.98	4.04*
Interaction: A x B	1	48.21	
Interaction: A x C	2	70.29	
Interaction: B x C	2	6.52	
Three-way interaction: A x B x C	2	61.36	
Error	53	57.86	

*p < .05

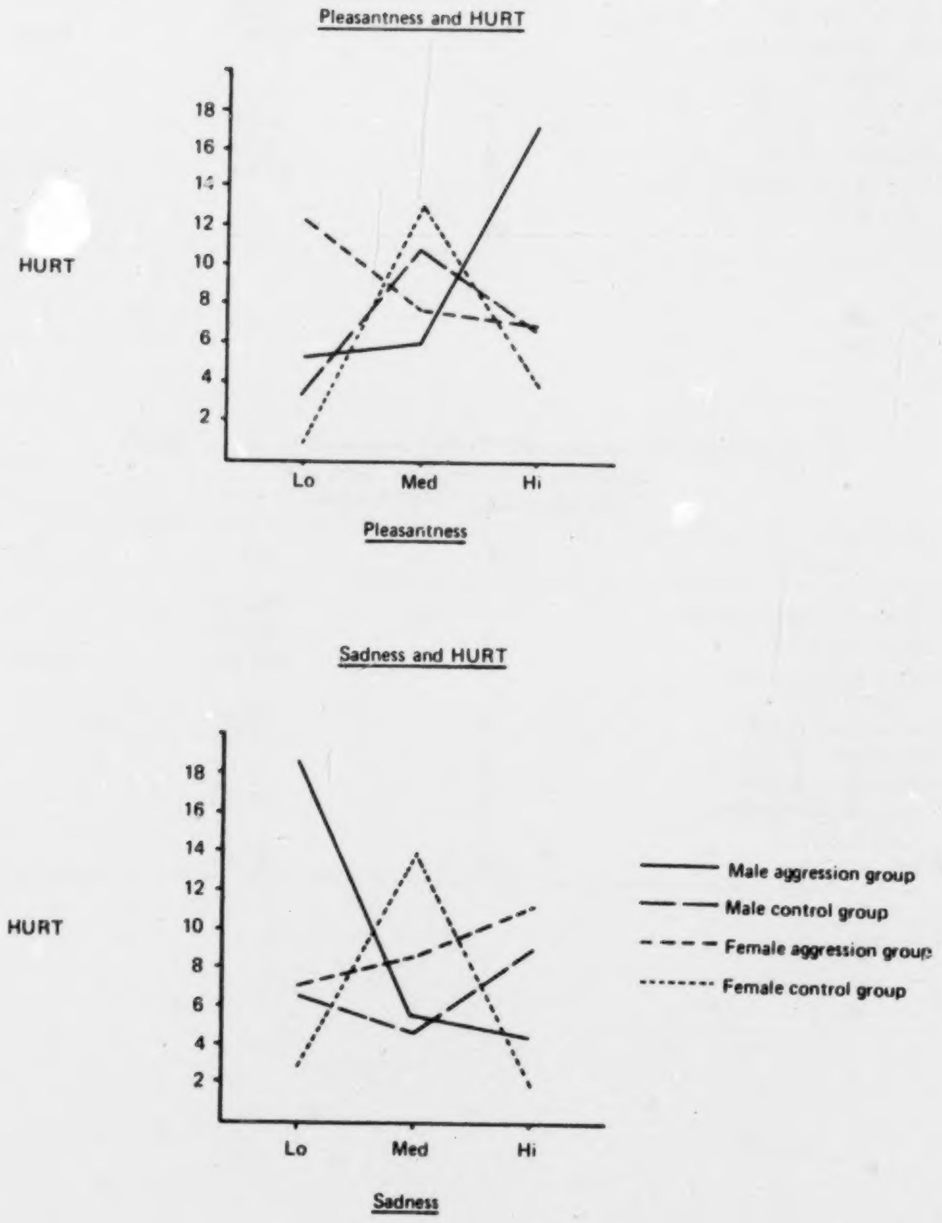

Pleasantness and HURT
HURT
18
16
14
12
10
8
6
4
2
Lo
Med
Hi
Pleasantness
Sadness and HURT
HURT
18
16
14
12
10
8
6
4
2
Lo
Med
Hi
Sadness
Male aggression group
Male control group
Female aggression group
Female control group

Figure 1

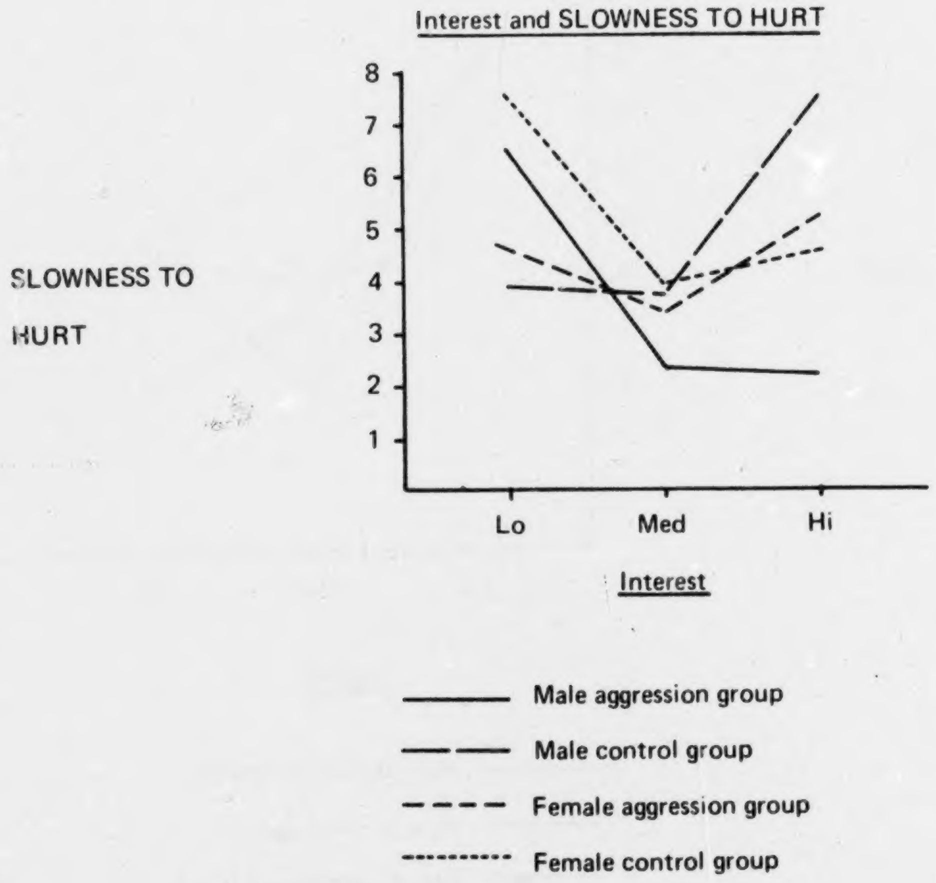


Figure 2: Interest and SLOWNESS TO HURT

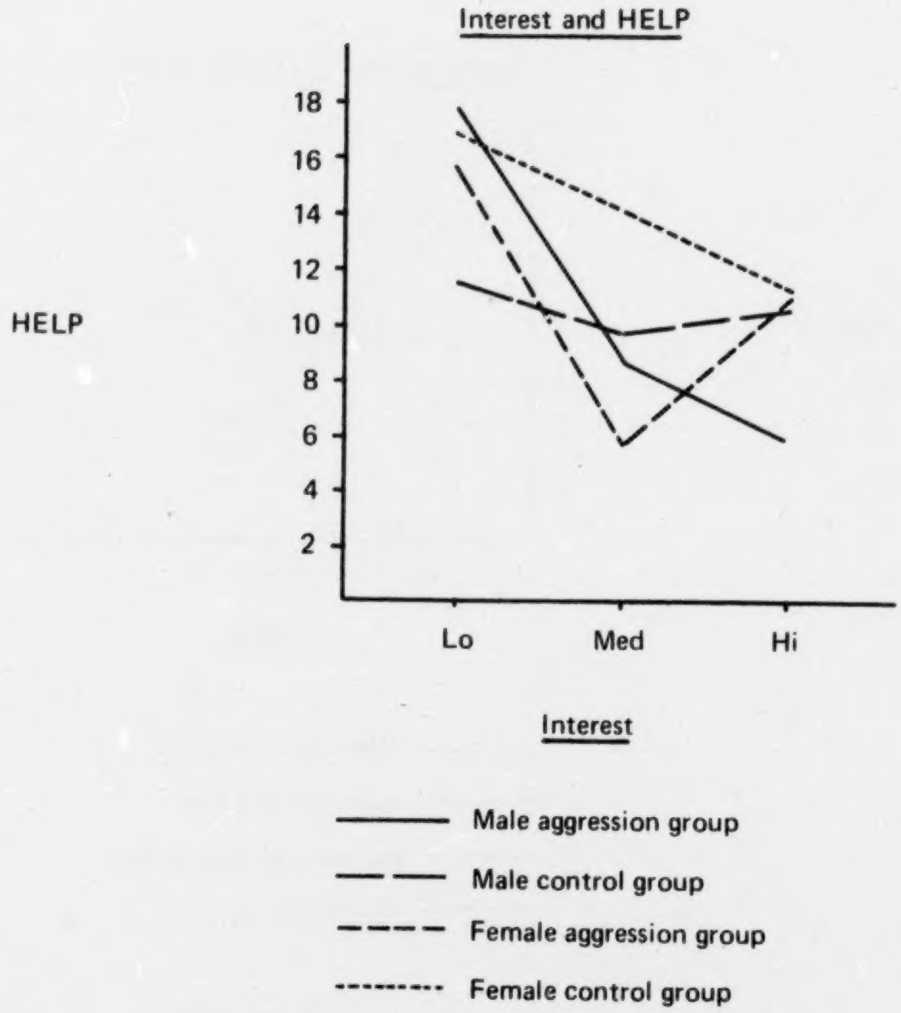


Figure 3: Interest and HELP

Televised Violence and Dream Content

David Foulkes, Edward Belvedere, and Terry Brubaker
University of Wyoming

A consensus has been growing among psychologists studying the effects of filmed fantasy violence that such filmed content stimulates the impulse to agress against others (Berkowitz, 1962, 1964, 1968; Bandura, 1965; Goranson, 1970; Siegel, 1969; Walters, 1966). These authors have not been averse to generalizing freely from a series of laboratory experiments conducted during the 1960s (Bandura, Ross, and Ross, 1963; Berkowitz and Rawlings, 1963; Walters and Llewellyn-Thomas, 1963). Thus Berkowitz (1964) concludes that "filmed violence is potentially dangerous. The motion picture aggression has increased the chance that an angry person, and possibly other people as well, will attack someone else" (p. 41). Discussing the alternative view that mass media violence may aid viewers in purging their own hostility, Walters (1966) concludes, "The highly consistent evidence from studies of aggressive models by Bandura, Berkowitz, Walters, and others suggests that the catharsis doctrine is not merely mistaken, but that its promulgation can lead to the defense of mass-media content that has socially harmful effects" (p. 590).

Such conclusions raise the question of how one extrapolates from the laboratory to the world outside the laboratory. In recent years, social psychologists have become increasingly aware of the fact that there often can be a great difference between having a seemingly "internally clean" experiment and being able to generalize with confidence beyond the boundaries of the conditions of that experiment (Campbell and Stanley, 1963). They have also become increasingly sensitized to the "game-playing" behavior in which persons will engage when they know that they are the objects of study and manipulation within a laboratory setting (Orne, 1962; Argyris, 1968). Thus the translation from laboratory to "real world" becomes a complex process, in which particular attention must be paid to details of experimental manipulations, subject instructions, and other factors that may jeopardize external validity.

It is not sufficient, for example, to argue, as do Goranson (1970) and Siegel (1969), that when results converge from several studies employing somewhat different techniques for manipulation and assessment of variables, one's confidence in their generalization to "real-world" social phenomena should thereby be enhanced. Were all these experiments to share common features biasing external validity, then no sheer total of results, however large, could justify the uncritical transposition of those results to contexts beyond the laboratory paradigms generating them.

Goranson does recognize that "it is important to know . . . just how far the results of these experimental studies can be applied to everyday media effects" (p. 28). He further notes that "we need additional research methods and experimental controls in natural, everyday media viewing situations." Similarly, Feshbach (1963) noted, "There is a very large gap to be bridged in generalizing from studies dealing with the impact of a single program to the possible effects of prolonged exposure to television programs with aggressive content" (p. 95). Interestingly enough, Feshbach's recent naturalistic investigation of relatively long-term effects of televised violence has produced results that contradict the trend of laboratory evidence; he found reductions of overt aggressive behavior, or no change, following exposure to a systematically selected "diet" of aggressive television programs (Feshbach and Singer, 1971).

The search for new laboratory paradigms

Thus Goranson's point that additional research methods are needed to evaluate the generalizability of filmed violence effects observed in the laboratory seems well taken. These methods need not be limited to application in naturalistic settings, however; even within the laboratory itself, new methods might be devised to attempt to surmount some of the difficulties inherent in previous attempts at generalization. To determine the form such new methods might take, it may be instructive to evaluate the designs of two studies from the gathering laboratory consensus whose generality is in question.

Representative of the many studies conducted by Berkowitz and associates is a study by Berkowitz and Rawlings (1963). Subjects were 160 college men and women. They were recruited from introductory psychology classes at the University of Wisconsin "to participate in a survey of reactions to movies and TV" (p. 406), and they received grade point credit for their participation. After their arrival at the laboratory, subjects were told by their experimenter that another graduate student would like to take advantage of their presence to collect normative data on an intelligence test. While giving the test, the graduate student either insulted the subjects or treated them neutrally. Then the original experimenter returned to show the subjects a brief prize fight scene from the movie *Champion*. Subjects were told that, after the viewing, they "would be asked to indicate their reactions to the depicted aggression" (p. 407). To "help them evaluate this scene," the experimenter introduced the film with statements in which aggression against the protagonist was portrayed either as justified or as unjustified. Following the conclusion of the movie, the experimenter indicated some concern that the subjects' reactions to the movie might have been affected by the prior intelligence testing and stated that she therefore would like to assess their attitudes toward that task. Included on a one-page questionnaire handed to subjects were two questions assessing their attitudes toward the intelligence examiner. The degree of punitiveness detected in response to these two questions was the primary dependent variable of the study. Questionnaire ratings supported the hypothesis that increased hostility would be "shown" toward the insulting graduate student following exposure to the film when the aggression portrayed therein had been described by the experimenter as justified.

As can be seen from this brief description, the Berkowitz and Rawlings study is representative of a form of social psychological experiment that has become increasingly popular in recent years (Kelman, 1967)—namely, the study in which independent and dependent variables are introduced by deceptive statements and acts of the experimenter. There is reason to believe that introductory psychology students, both by exposure to course content and by traditions passed along through the student culture, are familiar with the deceptiveness of many psychological experiments. As Argyris (1968) has argued, it seems likely that students will react to their manipulation in such experiments either by passively complying with what the experimenter seems to want or by retreating to private definitions of the experimental situation while feigning minimal interest in the task. In either event, the experiment has become a game with its own set of rules rather than a model for extra laboratory situations in which the person maintains his perceived freedom. Thus, control over variables within an experiment may be purchased at the cost of reducing the subjects' behavior to game-like strategies of extremely limited generality.

An often cited study by Bandura, Ross, and Ross (1963) may be taken as representative of research from Bandura's laboratory. Subjects were 48 boys and 48 girls from a nursery school at Stanford University. They were formed into four groups of 24 each. One group saw a "real-life" aggressive model; a second group saw the same actor engaging in aggression on film; a third group saw a cartoon-like character behaving aggressively; a fourth group served as a no-film control. Subjects in experimental and control groups had been matched individually on the basis of prior ratings of their aggressive behavior at the nursery school. The human model in the first two groups pummeled a Bobo doll in certain characteristic ways; in the third condition, a cartoon-like "cat" performed the same acts. After experimental-group subjects viewed a live or filmed performance, all subjects were mildly frustrated by being shown some attractive toys and being told that they could not play with those toys as the toys were reserved for other children. The children then were taken, in the presence of an adult, to another room in which a Bobo doll was present. During a 20-minute observation period in this room, their behavior was scored for directly and partly imitative responses of "aggression" to Bobo. Total aggression scores were higher in all experimental conditions than for the control subjects, with the human-live or human-filmed models eliciting the most direct imitations of the modeled aggressive behavior.

Several aspects of the design of the Bandura, Ross, and Ross experiment deserve comment. First, the stimulus was not a typical media stimulus, nor even—as in the Berkowitz and Rawlings study—a portion of such a stimulus. It was designed with the deliberate end in mind of teaching techniques of aggression, and it presented such techniques outside their usual dramatic context. While the subtlety with which preschoolers can perceive experimenters' intentions is probably quite limited, it is not inconceivable that these subjects could discriminate the intentionality behind their exposure to these experimental stimuli from other more impersonal and less directive contexts in which media stimuli are encountered. Second, the use of the postfilm frustration to enhance the expression of aggressivity certainly is justifiable within the confines of the experiment, but it does indicate potential limitations to the generality of results. Also limiting such generality is the unnaturally close link between the movie content and the posttest situation (a Bobo doll in each). Third, the experiment used a "response form" definition of aggression (Goranson, 1970), in which, for example, punching Bobo is an "imitative aggressive" response. However, punching a doll expressly constructed for that purpose is a response which allows for a multiplicity of mediating states, not all of which would meet a "harm-intent" definition of aggression. Diffuse arousal or excitement, for example, might lead to heightened activity in whatever modalities the testing environment allows; thus, if that environment includes a punchable doll, the latter will be punched, but perhaps more from an interaction of a heightened general drive level and situational constraints for its expression than from any kind of hostile intent.

Within a particular observational context, of course, a punch is a punch, no matter what its motivational mediation. However, differences in such mediation clearly are relevant to the question of the generalization of laboratory-induced punches to "real-life" socially aggressive behavior. To the extent that children in the Bandura et al. study punched the Bobo without intent to inflict harm, their behavior carries considerably less ominous implications for the real world than if they were venting film-instigated hostile feelings.

The preference of recent investigators for "response-form" or simple behavioristic definitions of aggression has had other unfortunate consequences, as we may see when we refer back to the Berkowitz and Rawlings study. These investigators wished to refute Feshbach's (1961) cathartic interpretation of his finding that insulted subjects shown an aggressive film subsequently indicated less hostility toward the experimenter than did insulted subjects shown a nonaggressive film. Berkowitz and Rawlings suggested that the decreased expression of aggression was due to an increase in guilt-related inhibitions in the angered subjects shown the violent film. Their justification manipulation was an attempt to test this alternative interpretation, the idea being that justified aggression would be less likely to induce guilt over the subjects' own hostility than would unjustified aggression. The results were in accord with their prediction. That the justification-nonjustification manipulation worked through the mediation of differential guilt arousal, however, is an assumption only indirectly suggested by these authors' manipulations and results. Choosing relatively simple behavioral measures as their dependent variables, these authors did not attempt more direct assessment of the dynamic, covert processes that may have guided their subjects' behavior.

The Berkowitz and Rawlings and the Bandura, Ross, and Ross experiments have been considered here in some detail because they generally are thought of as relatively strong links in the chain of evidence leading to the conclusion that media violence stimulates (rather than reducing or having no effect on) the expression of everyday aggression. The intent of the critique is not to dispute the findings of the two studies in question or of the research programs of which they are representative. Rather it is hoped that the discussion might indicate several respects in which the generality of current laboratory studies is questionable and suggest alternative manipulations and dependent variables for future laboratory research to clarify this question of generality. With these goals in mind, what procedural recommendations might be made for future research on the effects of televised violence?

Limitations of current trends in laboratory experimentation. Goranson argues:

> While the laboratory situation may not represent in detail the ordinary movie theater or living room, certain important aspects of the real-life setting may be represented in the laboratory—the television set or film screen, the audience, etc. Other features of the living room, such as table lamps or drapes on the wall, for example, may be quite unimportant in determining whether the findings obtained in the laboratory will hold for the American home. In other words, it is not really necessary to duplicate the home or theater in complete detail in order to have *some* degree of generalization from the laboratory to the real world.

In focusing on physical similarities and dissimilarities of laboratory and home, this argument is quite misleading. The important limitation of contemporary laboratory research in social psychology lies in the fact that the laboratory becomes the setting for a unique form of *social interaction*, one with its own peculiarities of role behavior for all parties concerned. It is a place where experimenters deceive subjects and where the latter are often aware of it but must try as best they can to comply with the apparent demands of the experimental design without either getting too involved or revealing too much of their everyday selves.

The discrepancy between the results of Feshbach and Singer's (1971) naturalistic experiment and earlier laboratory studies might be taken to indicate the desirability of employing more naturalistic procedures within the laboratory itself. If the laboratory need not be made more physically similar to usual viewing settings, then the subject must be made to feel more "at home" there. He must be treated more as he is treated in everyday social interaction (Kelman, 1967). While deception and manipulativeness cannot be eschewed, they also should not be endlessly elaborated without regard for the undesirably limited social situation thereby created.

Subject selection. One means of circumventing the compliancies and counterstrategies of suspicious subjects, then, is for the experimenter to behave less suspiciously. But for the current crop of college students, that device may be too late and too little reinforced by other encounters with the behavioral sciences. In conjunction with less contrived experiments, however, one might well consider the use of subject populations other than college students (Schultz, 1969). For media-violence studies, interest in generalizing to salient social concerns might dictate using children as subjects. Preadolescents would seem to be ideally suited for this purpose. They are old enough and well enough schooled (in the literal sense) to understand the importance of accuracy and motivation in responding to experimental tasks, yet young and inexperienced enough not to behave as reactively (Campbell and Stanley, 1963) as college students or adolescents do. (For example, they seem less likely in a relatively uncontrived situation to modify their responses in accord with their perceptions of the experimenter's goal.) The preadolescent, moreover, is, in contradistinction to Bandura et al.'s (1963) preschoolers, at an age where stimulus effects are likely to be of immediate *social* relevance, and, in contradistinction to Walters and Llewellyn-Thomas's (1963) mid-adolescents and Berkowitz and Rawlings's (1963) college students, he is at an age at which consumption of the media violence of "adult" programs is a normative behavior. (Schramm, Lyle, and Parker [1961] report television viewing to be heaviest during grades six, seven, and eight.)

Use of typical mass media stimuli. Children usually are not exposed to a brief segment of a dramatic movie taken out of its context, nor to "training films" on how to pummel a Bobo doll. The use of such stimuli limits generalization to real-life social contexts in at least two ways: it lessens stimulus similarity, and thus stimulus generalization, and it probably increases subjects' perceptions that the experimental situation is an unnatural one and hence one to be reacted to in a qualitatively different fashion from real-life viewing situations. It seems desirable to supplement experimental designs in which atypical stimuli are required for purposes of internal control with designs in which typical stimuli are employed to assess external relevance.

Reactivity and situational specificity of dependent variables. Generalizability of findings from laboratory experiments to real-life situations is enhanced when posttreatment measures do not bear an obvious relationship to the experimental treatment (Campbell and Stanley, 1963). On the other hand, when the relationship between the two is close, the posttest gives rise to "clues in divining the experimenter's intent. . .[and] play-acting, outguessing, up-for-inspection, I'm-a-guinea-pig. . .attitudes" (Campbell and Stanley, p. 190). Goranson (1970) concludes that similarity of violent media content to postexposure test conditions enhances the demonstration of increased aggressiveness; it may merely show that subjects in the face of such similarity are better able to comply with the implicit demands the experimenter has set for them. Not only is seeing a Bobo doll in the posttest situation atypical of real-life contingencies; it also signifies artificial restraints on the range of typical postexposure responses. Treatment-posttest similarity also may generate the idea that the experiment is a qualitatively different situation than one faces in real life, one to be responded to accordingly (unnaturally). The generality of laboratory findings on the effects of mass media violence will become more clear, then, when measures of postexposure aggressiveness are: (a) somewhat removed from the specific content of the media presentation; (b) relatively unbounded by any unnatural restraints on free responding imposed by conditions in the posttest; and (c) relatively insusceptible to voluntary control in the service of goals of acquiescing to perceived experimental demands.

Overt behavior and inferred motives as dependent variables. Typically, those studies which constitute the "growing consensus" of media researchers that filmed fantasy violence is harmful have employed behavioristic reasoning in selecting their measures of aggression. Thus, dependent variables often are checkmarks on rating scales (Berkowitz and Rawlings, 1963) or overt motor responses of a certain form like hitting Bobo (Bandura et al., 1963). These conventions have kept the dependent variable simple and its assessment reliable, but they have not made for corresponding unequivocality at the interpretive level. We have argued that questions of generality particularly require knowledge not only that a certain "aggressive response" was given but also of its motivational mediation. Even within the behavioristic tradition of media stimuli cuing aggressive responses, recourse has been taken in complex mediating mechanisms such as increased guilt, but researchers have not devised measurement techniques commensurate with these theoretical developments.

It has often been stated that media entertainment content consists of more than first meets the eye or ear, that the manifest content of the program overlays a latent, symbolic message (Barker, 1955; Emery, 1959; Homans, 1961). Feshbach and Singer (1971) state, "Clearly, whatever its specific effects, the depiction of human violence produces results through stimulation and interaction with the individual's fantasy life" (p. 4). If this is in fact the case, then immediate and long-range effects must be sought at a level where latent and symbolic fantasy processes are operative. Since Freud's *Interpretation of Dreams* (1900), it has been assumed, with good supporting clinical data, that the dream is a prime example of such a level of psychic functioning. In dreams the course of mentation is guided less by external, social stimuli and more by endogenous, intrapsychic stimuli. In dreams mental associations become less bounded by the rules of waking logic and mental content, freed from a process of social monitoring, can deviate toward the impulsive and relatively nonsocialized dimensions of personal functioning. It is precisely at this level that the dependent variables of traditional studies of the effects of media violence seem most limited.

Where the motives underlying waking responses are in question, dream data might be an especially helpful supplement to observations of overt, waking behavior (Foulkes, 1966). When guilt and anxiety about aggressive impulses are postulated as critical variables mediating overt behavioral expression of aggression, it would seem relatively easier to assess such variables in covert fantasy than in motor behavior. Dream content, moreover, meets the several other criteria we proposed for supplementary dependent variables in research on the effects of media violence: it is generated under conditions in which experimentally imposed restraints on response freedom are relatively lacking (in a posttest play situation, one can play only with those toys the experimenter provides; one can dream literally of anything); it is the product of relatively involuntary processes and is remarkably resistant to substantial, direct control by presleep or sleep manipulations (Foulkes, 1966); and, in light of this manifest lack of connection with particular waking experiences under experimental manipulation, it fulfills the criterion of stimulus-posttest distance.

The present paper reports a study in which dream content was employed to assess the effects for male preadolescents of a typical media portrayal of fantasy violence.[1] To judge by the record of past research, dream content has not seemed to many media researchers to be a fruitful place to look for such effects. Our introductory survey of what *has* been capturing the attention of media researchers, however, tried to indicate that current practices in experimental design and variable selection in this area have not been entirely adequate, that supplementation of these practices with new and different research techniques is desirable, and that a variable like dream content has something of value to offer the researcher in this regard. Before we embark upon a description of our implementation of this line of reasoning, however, it will be necessary to say something more about dreams, in order that the feasibility of their study and their relevance to the topic at hand—the effects of fantasy violence in the mass media—may be clarified.

Recent studies of dreaming

Historically, the use of dream content as a dependent variable in social psychological research encountered two major obstacles, one methodological and the other conceptual. Technology was lacking for representative sampling and for providing conditions of reasonably faithful reporting of sleep mentation; no empirically well-supported conceptual framework was available to link dreams with waking dependent variables of more immediate interest. In recent years, however, each of these obstacles has been surmounted, and studies of dream content might now routinely be considered in any comprehensive program on the effects of significant social stimuli.

Beginning with the studies of Aserinsky, Dement, and Kleitman (Aserinsky and Kleitman, 1955; Dement and Kleitman, 1957), it has been demonstrated that periodically recurring episodes of rapid eye movement (REM) sleep are associated with vivid dreaming and that adult subjects awakened from such periods are able to report dreams in considerable detail on 80-90 percent of experimental trials. As discrete periods of REM sleep occur three to six times nightly for all subjects (consuming as much as 25 percent of total sleep time), the REM-awakening technique of dream retrieval is capable of producing considerably more dream material, more representatively sampled dreams, and dreams reported considerably closer to their time of occurrence than are older daytime-questionnaire techniques of dream elicitation. Since the Aserinsky-Dement-Kleitman techniques were first developed, the retrieval of dream content from REM sleep has increasingly been exploited by researchers to answer a variety of questions on the psychophysiological and experiential determinants and correlates of dream experience (see, for example, Foulkes, 1966).

Research in this vein has contributed to an evolving framework relating dream content to waking experience. Classical dream theory had been of several minds with respect to the direction of waking-sleeping correlations. Some (Jung, 1948) had argued for the complementarity of sleeping and waking experience—hence the hypothesis that the waking introvert would demonstrate extroversion in his dreams. Others (Adler, 1931) had felt that dreaming and waking were essentially continuous, exhibiting the same characteristic concerns and styles of dealing with reality. In recent years many REM-retrieval dream content studies have been conducted which are germane to the issue of complementarity *vs* continuity. It has been shown, for example—contrary to Jung's specific prediction—that boys who score high on a waking measure of social dominance (California Psychological Inventory) tend to report REM dreams in which they assume an active role, while those low in waking social dominance play a more passive role in their reported REM dreams (Foulkes, Larson, Swanson, and Rardin, 1969).

A number of other studies point in the same direction. For instance, subjects with greater demonstrated pathology in wakefulness (as indicated by Minnesota Multiphasic Personality Inventory profiles or by the fact of institutionalization for emotional disturbance) have more bizarre REM-sleep mentation (Foulkes and Rechtschaffen, 1964; Pivik and Foulkes, 1966; Foulkes et al., 1969). More significantly for research concerning drive variables, Ben Horin (1967) found that high expressers of hostile impulses in wakefulness (as assessed by the Minnesota Multiphasic Personality Inventory Ho scale and by ratings of overt behavior) were high expressers of overt hostility in dream content. (Interestingly enough, *waking* fantasy [Thematic Apperception Test] measures of hostility related consistently neither to more overt waking variables nor to dream content ratings.) Thus dream content ordinarily seems to reflect chronic personality characteristics, including those relating to impulse expression, sufficiently well to be employed as a supplementary measure of these waking variables.

Laboratory studies of presleep manipulations of drive variables also lend some support to the hypothesis of continuity between short-run fluctuations of drive states in wakefulness and the expression of related drive content in dreams. Hauri (1966), for example, has shown that probable *satiation* (through extensive presleep physical or mental exercise) of drives to physical or mental activity led to dreams with *reduced* motility or intellective content. On the other hand, when deliberate efforts have been made to ensure *deprivation* of waking drive activity, *increases* in drive-related content have been observed in dreams. Thus Wood (1962) found that presleep social deprivation led to an increase in socially active mentation during REM sleep. Likewise Whitman, Pierce, and Maas (1960) found that a subject given a muscle relaxant (Meprobamate) had dreams containing increased physical motility content. Possibly related findings are those by Swanson and Foulkes (1968) that college coeds' dreams were most manifestly sexual during that phase of the menstrual cycle associated with low waking sexual interest (and expression?) and those of Witkin and Lewis (1965) that subjects who viewed traumatizing sexual films (showing mutilation of the sexual organs) before going to bed, and who could not be presumed to have mastered the anxiety thereby created, had frequent REM dreams with obvious sexual symbolism related to particular elements of the film.

In general, then, drive expression observed in dream content seems to be positively related both to chronic levels of such expression in wakefulness and to short-run fluctuations in hypothetical drive states as reflected in presleep experiences of probable satiation (reducing effective drive state and subsequent dream expression) or deprivation-stimulation (increasing effective drive state and subsequent dream expression).

Continuities in mood states between wakefulness and sleep have also been observed. Hauri (1966) found a continuity of tension from a presleep study period into subsequently retrieved non-REM dreams. Swanson and Foulkes (1968) found that college girls' REM dreams were most unpleasant during that phase of the menstrual cycle in which they reported experiencing the greatest waking depression.

Considering recent media research has led us to the conclusion that new dependent variables are required which are less reactive, more remotely connected with the stimulus whose effect is being assessed, and more reflective of the person and less reflective of a particular assessment situation than those variables that have hitherto been employed in this area. We have suggested that dream content variables meet these specifications. This suggestion, however, might raise a new question—namely, whether dream content is so far removed from the overt behavior of traditional concern to media researchers as to carry little meaning outside its own limited domain. Dream variables do, of course, have some intrinsic interest. Increased incidence of anxiety dreams after exposure to violent films, for example, could be considered a harmful effect in its own right, and there is evidence that some children at least think that violent television programs "give you bad dreams" (Himmelweit, Oppenheim, and Vince, 1958). Yet the evidence just reviewed suggests that dreams are more than symptoms of isolated interest. The world of REM-sleep mentation to which ready access has recently been gained does not seem to be part of another universe of experience, foreign to wakefulness. Rather, it seems an integral component of a unitary experiential world comprising both wakefulness and sleep. Perhaps for the first time, the student of dreams is in a position to supply a coherent, experimentally supported framework with which dream observations may be linked to waking variables of more immediate social relevance.

Filmed violence and dreams: prior research

Although not cited in various reviews of the literature on the effects of media fantasy violence, two studies employing dream-content dependent variables had been performed during the 1960s. Foulkes and Rechtschaffen (1964) studied the effects of an immediately presleep exposure to a television film upon dreams subsequently reported on awakenings from REM and non-REM sleep. (Non-REM sleep is associated with some memorable sleep mentation which is, however, relatively less dreamlike than the mentation of REM sleep—Foulkes, 1962; Rechtschaffen, Verdone, and Wheaton, 1963; Monroe, Rechtschaffen, Foulkes, and Jensen, 1965). The purposes of this study were several; one was investigating the contribution of day residues to nocturnal dream content. The study also was conceptualized as investigating a new and possibly more subtle dependent variable for assessing the effects of mass media entertainment content—one capable of revealing affective and cognitive reactions at a symbolic and relatively involuntary level of free organismic responding.

The stimuli in this 1964 study were two episodes of a television western series, exposed singly and in counterbalanced order on each of two nonconsecutive nights. One was a typical episode of the series, in which the hero tracked down two robbers. Violence was abundant. The other episode was atypical of the series as a whole. In it, the same hero served as an unwilling matchmaker. This episode was essentially a light, romantic comedy which happened to be set in the Old West. Subjects were 24 paid volunteers, mostly in their twenties, and generally either university students or employees. Thirteen were male and 11 were female. Dream reports were solicited on a scheduled two non-REM and three REM awakenings per night (with a pooled group recall of 88.8 percent on the 143 REM, and 61.9 percent on the 84 non-REM, awakenings actually effected). Dreams were analyzed in several ways: subjects completed rating scales on their dreams' clarity, distortion, emotionality, hedonic tone, on the degree of control exercised over dream events, violence-hostility, degree of activity of the self-character, and dramatic quality; two judges rated the dreams blind for their imaginativeness; two judges, one blind, rated the dreams for verbal aggression need and press, physical aggression need and press, and heterosexuality need and press; and two judges, one blind, attempted to postdict the film seen on the basis of apparent incorporation of film material into dream reports. Reliabilities for various judge-rated variables generally were satisfactorally high.

Effects of the film manipulation may be summarized as follows: no difference in latency to sleep onset or initial REM period; no difference in either REM or non-REM dream recall rate; significantly longer REM reports following the violent episode; significantly more imaginative REM reports following the violent episode; REM-sleep dream reports rated by subjects as significantly more vivid and emotional following the violent episode; non-REM dream reports rated as significantly more distorted following the violent episode; no significant effects on any need or press dream rating variable; very low levels of direct stimulus incorporation (five percent of 179 dreams showing unequivocal incorporation of a film element), with no significant difference between the two films in the incidence of such incorporations. It was concluded that the effect of the violent episode was to produce REM dreams that were more exciting and interesting in general without influencing their specific aggressive content or hedonic tone.

Foulkes and Rechtschaffen noted some qualifications of the possible generality of the results of their study: (1) The film stimuli differed in other respects than in their specific drive content. The violent film was less perceptually and thematically differentiated than the nonviolent one. The films also differed in perceived artistic merit; 16 of 22 subjects who saw some difference between the two films thought the nonviolent film the better of the two. In any study using a film "stimulus," there is the problem of identifying just which elements of the stimulus complex actually are in some functional relationship with the dependent variables. (2) Subjects in the study were "intellectual" young adults who, from the evidence of their own postfilm questionnaires and of unsystematic observation by the projectionist, generally were uninvolved in the film-viewing situation. Discussing this aspect of the study, the authors specifically raised "the question of what kinds of effects might be observed with less self-conscious subjects who might become highly involved during the showing of such films" (p. 999).

Such speculation served as the bridge to a second study of the dream effects of viewing violent fantasy (Foulkes, Pivik, Steadman, Spear, and Symonds, 1967). This was the first systematic investigation of children's dreams using the REM monitoring technique, and the published presentation of its results tended to focus on normative aspects of children's dreams. Nevertheless, the study originally was conceptualized as an extension of the earlier film study with young adults and did include a systematic presleep film-stimulus manipulation. The stimuli were two ten-minute films available for home distribution: *The Indian Raiders,* a manifestly aggressive western, and *This is Baseball,* a documentary on Little League baseball with passing references to major league baseball.

Subjects were 32 boys, 16 aged 6-8 and 16 aged 10-12, who slept for two nonconsecutive nights in a sleep laboratory. Eight of the younger and eight of the older subjects viewed the films in the order night 1—Western, night 2—Baseball, and the remaining subjects viewed the films in the reverse order. Films were shown immediately before the subjects were put to bed. Dream reports were solicited on a scheduled four REM-sleep awakenings per night. Overall, subjects recalled dreams on 179 of 249 awakenings actually effected, for a recall percentage of 72 percent. Dream dependent variables included: content analysis categories in the areas of plot, setting, and characters; ratings by naive judges of dream distortion, hedonic tone, activity of the subject's own role in his dream, hostility need, hostility press, imagination, and so on (interjudge reliability for these particular ratings was satisfactorally high except for hostility need, which was very rarely observed); substantive word counts; blind judgments by two observers of manifest film incorporations; and subjects' own judgments as to whether their dreams were "good" or "bad."

Effects of the film manipulation were as follows: no difference in latency to sleep onset or initial REM period; significantly less judge-rated imagination, goodness of recall, and hostility press following the violent film; significantly fewer subject-classified "bad" dreams following the violent film; fewer "Social: Escape/Hostile" classifications by the content analyst of dreams collected following the violent film. It is clear that the pattern of results here is quite a contrast with those observed by Foulkes and Rechtschaffen in young adult subjects. In the earlier study, the results pointed to *increased nonspecific activation* following the viewing of a violent film; in the later study, they indicated *reduced general activation and reduced levels of* the *specific drive content* contained in that film. The latter results seemed strikingly inconsistent with Berkowitz's (1962) general position on the instigation of aggression by violent films.

The major conceptual problem at this point became the reconciliation of the children's results with those of the earlier study. The subjects differed in age, of course, but age *per se* is not a psychological-process variable capable of accounting for the discrepancy. One obvious difference between the two studies was in the degree of involvement subjects brought to the task of film viewing. In the child study, the projectionist rated each subject's manifest interest in the stimulus on a five-point rating scale (1 = quite inattentive, to 5 = rapt attention). Interest in the films was quite high, with an overall mean rating of 4.27. Significantly greater involvement (4.47 vs 4.06, $p < .01$) was observed for the violent film, although the baseball film was of absolutely high interest value. Qualitative observations by the projectionists of the incidence of such activities as nervous fingering of EEG electrodes or genitalia also indicated especially high involvement during the viewing of the violent film. The authors suggested that they had elicited a sufficient degree of involvement during the viewing of the aggressive stimulus to permit catharsis. As Schramm et al. (1961) have observed, "It is unlikely that television will either reduce or increase tension unless it provides something with which the child can identify. To the extent that he stands aside and 'observes' the program, to the extent that he employs . . . 'adult discount', he is unlikely to get much emotional experience out of the program. Therefore, unless he can go through some of the aggressive experiences of his Western hero . . ., there is not likely to be much change in his state of tension" (p. 133).

Some data were collected relevant to Berkowitz and Rawlings's (1963) alternative hypothesis accounting for postexposure reductions in hostility in terms of increased guilt. A six-point (0-5) "guilt" scale was employed to assess a dreamer's degree of acceptance of personal responsibility for hostile actions, primarily expressions initiated by others in his dreams. At low scale points, hostile press was absent or, if present, subject to no justification; at middle scale points, a rationale was presented for hostile press wholly or primarily in terms of others' motives; at high scale points, responsibility for hostile press was accepted primarily or wholly by the dreamer himself. Reliability of these ratings was quite low (.48), at least in part because so little dream aggression was felt by subjects to require any explanation or personal acceptance of responsibility. The mean rating subjects received on the scale was .35. This in itself might be used to refute Berkowitz and Rawlings: subjects showed so little apparent guilt over the hostile impulsiveness they ascribed to characters in their dreams that it seems unlikely that such guilt mediated the overt hostility outcomes in dreams collected following exposure to the two films. Moreover, in spite of low incidence and low interrater reliability, *less* guilt was scored in dreams collected following the violent stimulus than following the baseball control film (.28 *vs* .41, marginally significant at $p < .10$).

Another finding linked high involvement in viewing a violent film with decreased dream aggression. The high absolute attention shown by all boys to both films placed statistical limitations on the correlation of individual film reactions to subsequently retrieved dream content. However, in support of the catharsis hypothesis, a nonsignificant trend ($r = .22$) was noted for boys with the most dream hostility press following viewing of the nonviolent film relative to the amount scored following the violent film to have been the most involved during the presentation of the violent film. Considering both the young adult and the preadolescent studies, it is also interesting to note that in each case the stimulus with more immediate interest value for subjects (the romantic comedy of Foulkes and Rechtschaffen [1964] and "The Indian Raiders" of Foulkes et al. [1967]) was associated with less intense subsequent dreams.

Filmed violence and dreams: directions for further research

Each of the two studies employing violent media fantasy as an independent variable and dream content as a dependent variable found some consistent intrastudy effect of the former upon the latter. This seemed a promising indication that dreams *can* be used to assess the impact of media stimuli. That the effects were not consistent across the studies, however, indicated the desirability of further research in which the contribution of different film, subject, and viewing-context variables to overall dream effects could be assessed. Such research not only might clarify the discrepancies between the two earlier studies but also could indicate more generally the conditions under which certain effects will or will not occur.

Two related variables—subject's history of interest in the stimuli to be viewed and his involvement during the act of viewing—were suggested above as possible mediators of different dream effects in the study of young adults (Foulkes and Rechtschaffen, 1964) and the study of male preadolescents (Foulkes et al., 1967). This *posthoc* explanation seemed deserving of an independent experimental test. In the present experiment, we have attempted to provide such a test.

Prior exposure rate. This subject-selection variable was suggested by differences in age level and presumed contemporary television behavior between subjects in the two earlier film-dream studies. Much evidence has indicated that children's habitual exposure rate to television in general and to violent media fantasy in particular is associated with a number of other variables that are pertinent to concerns about the social effects of everyday television viewing. Among these variables are degree of parental restriction over media exposure (Bailyn, 1959), personal and social adjustment (Maccoby, 1954; Riley and Riley, 1954; Himmelweit et al., 1958; Bailyn, 1959; Schramm et al., 1961; Feshbach and Singer, 1971), intelligence and social class (Himmelweit et al., 1958; Bailyn, 1959; Schramm et al., 1961), and the child's aggressiveness or inhibitions over aggressive expression (Schramm et al., 1961; Eron, 1963; Feshbach and Singer, 1971).

Although a high viewing rate for televised violence cannot be totally independent of the overall rate of viewing television, and although prior research has indicated that the two viewing variables may have different correlates and meanings at a given age level, a synthesis of the results of such research suggests that the following characteristics might also discriminate our high-exposure subgroup from the low-exposure subgroup: lesser *vs* greater parental restriction on children's viewing behavior; lesser *vs* greater personal and social adjustment; lower *vs* higher intelligence; lower *vs* higher social class; and greater *vs* lesser waking aggressiveness. We will report some data on the status of our sample subgroups on most of these variables. Prior research also suggests, however, that no very substantial correlates of high *vs* low exposure to televised violence might be expected. With the possible exceptions of intelligence and social class, none of the variables heretofore investigated has convincingly and consistently emerged as a probable correlate of high exposure to, or stated preference for, media fantasy in general or aggressive fantasy in particular. Also, as Schramm et al. (1961) have observed, all 10-12-year-old children tend to be relatively heavy viewers of television, primarily of "adult" (i.e., often violent) programs; the psychological and social differentiation of high- and low-exposure children becomes most prominent only later in adolescence when children of certain characteristics (higher intelligence, better adjustment, higher social class) begin to reduce their exposure. Whatever personal and social variables might be associated with differential exposure to fantasy violence at any age level, however, the exposure variable remains of intrinsic interest to researchers speculating about the antecendents of "addiction" to fantasy violence and about the long-term effects of viewing media violence.

Involvement. The involvement variable in our study was suggested by differences in qualitative observations of viewing behavior and in effects between studies by Foulkes and Rechtschaffen (1964) and Foulkes et al. (1967). The act of viewing itself has been subject to little systematic study in the numerous investigations of the antecedents and consequences of such viewing. How a child watches fantasy violence has been of less interest to media researchers than why he watches it or what difference it makes that he watches it. Some studies of physiological processes during film viewing have been made (Dysinger and Ruckmick, 1933; Lazarus, Speisman, Mordkoff, and Davison, 1962), but generally they have been more concerned with the antecedents of differential physiological reactivity during film viewing than with the short- or long-run consequences of such differences in behavioral domains related either to the film stimulus or to the emotional reactions it produced.

From a theoretical point of view, it might seem that any film effect would be enhanced by focal, as compared to incidental, involvement during the viewing process. The cathartic position—that elicitation and vicarious expression of hostile impulses during the viewing of an aggressive fantasy reduces subsequent aggressive behavior—strongly implies that relatively great involvement and identification with the fantasy content are preconditions for the cathartic outcome. Schramm et al. (1961) already have been quoted to this effect. It might also be expected that any theory stating that aggressive fantasy stimulates aggressivity would imply that the effect would increase as a function of subject attention to the fantasy stimulus. Here, however, another alternative is suggested by some research findings (Pine, 1960; Allyn and Festinger, 1961) indicating that stimuli poorly attended to (exerting their effects on the "margin of consciousness") may have effects that stimuli focally attended to do not have. One explanation for such results has been that a subject's defenses are "down" when he is not focally aware of a stimulus. On balance, however, it seems more likely to predict that for drive-stimulation outcomes, as for drive-reduction outcomes, the effect will be proportional to a subject's attentiveness and involvement in the fantasy stimulus.

METHOD

Design

The design of the present study was a 2^3 factorial design with incomplete blocking. The first factor was a subject variable of high *vs* low prior exposure to televised violence; the second factor was a film variable (violent *vs* nonviolent); the third factor was a viewing-involvement variable (focal *vs* incidental). The assessment and manipulation of these three independent variables will be described below. Here we present a summary of the various groupings of temporally-ordered treatments to which high- and low-exposure subjects were randomly assigned:

Experimental Night 1:	Experimental Night 2:	Subject Nos:	
Film and Viewing Condition	Film and Viewing Condition	High Exposure	Low Exposure
Violent Focal	Nonviolent Focal	3	3
Nonviolent Focal	Violent Focal	2	2
Violent Focal	Nonviolent Incidental	2	2
Nonviolent Focal	Violent Incidental	3	3
Violent Incidental	Nonviolent Incidental	3	3
Nonviolent Incidental	Violent Incidental	2	2
Violent Incidental	Nonviolent Focal	2	2
Nonviolent Incidental	Violent Focal	3	3
		20	20

Thus all 40 subjects saw both films, 20 in the order violent-nonviolent and 20 in the order nonviolent-violent. Ten subjects viewed both films under focal involvement, ten viewed both under incidental involvement, and 20 saw one film under focal involvement and one film under incidental involvement (ten subjects in the order focal-incidental and ten in the order incidental-focal). The analysis of variance procedure employed allows the direct derivation of all interaction terms and of main effects for film and involvement. The main effects of addiction were calculated separately with t-tests for independent samples.[2]

Major interest attached to the following questions about dream dependent variables:

1. Does presleep viewing of a violent film reduce (or increase) general dream intensity, unpleasantness, and specifically aggressive content?
2. Does viewing involvement interact with reactions to the films, so that, for example, high involvement to a violent film decreases vivid and aggressive dream content ("catharsis") while low involvement does not affect, or enhances, such content?

3. Does the presleep viewing of a violent film have different effects for subjects who have had high and low prior exposure to such stimuli?
4. Are the dreams of high-exposure subjects more hostile, vivid, or distorted than those of low-exposure subjects?
5. Will increased intensity of viewing involvement, regardless of what is being watched, systematically decrease (or increase) the intensity or vividness of subsequently retrieved dreams?

The first three comparisons relate to main and interactive effects of exposure to films differing in manifest violence, and they are the major focus of the study. The other two are main effects for the remaining independent variables. The comparison in Question 4 is of interest in relation to speculations either about the long-term *effects* of viewing televised violence or about the *antecedent* conditions which lead subjects to different habitual exposure rates; the comparison in Question 5 offers a test of the idea advanced above (as a *post hoc* reconciliation of results of two earlier studies [Foulkes and Rechtschaffen, 1964; Foulkes et al., 1967]) that high presleep viewing interest leads to reduced levels of dream vividness.

Subjects

Recruitment. An advertisement in the Laramie daily newspaper brought 105 10-12-year-old boys to the University of Wyoming Sleep Laboratory to take a screening test designed to measure the exposure variable. The advertisement stressed the financial advantages—$25 payment for sleeping three nights in the laboratory—of participating in the experiment, while also noting that service in the study was "an interesting—and relatively effortless—way for boys to earn a generous sum of money while also helping University scientists in their studies of patterns of sleep and dreaming in children."

The screening test attempted to disguise the basis of subject selection by including a number of items relating to sleep behavior, demographic variables, and other leisure activities besides television viewing. Pages 6-9 of the questionnaire included a listing of regularly scheduled television entertainment programs appearing on three Denver network-affiliated stations during the week of June 13-19, 1971, between the hours of 4 and 10 p.m.[3] Subjects were asked to indicate, for each program listed, whether they viewed it "every week," "most of the time," "sometimes," "hardly at all," or "never." The actual basis for subject selection was the number of the 18 programs judged to be violent that the potential subject reported watching "every week" or "most of the time." two groups of 20 subjects each were formed on the basis of highest and lowest reported exposure to these shows. The 18 "violent" programs included those 15 which Greenberg and Gordon (1971) found that a sample of citizens of the Detroit area felt were the most violent programs on network television at the time. (Two other current network programs had not been rated by the Detroit panel, and the eighteenth program on our list was a syndicated rerun of a series from an earlier season). Among the subjects chosen for high exposure to television violence, the median number of the 18 shows watched very frequently was nine (range 7-17), while for the subjects selected for low exposure to television violence, the median was one, and the range was zero to two.

Other characteristics. From other responses to the screening questionnaire, it is possible to characterize the two subgroups (high and low exposure to televised violence) in terms of several variables which prior research or speculation has suggested may be correlated with the selection variable: viewing rate and preference for media fantasy in general; social class; intelligence; degree of integration into peer society; relationship with parents and siblings; symptoms of psychological disturbance; degree of parental restriction on television viewing; and parental encouragement of aggressive behavior.

The groups of subjects high and low in exposure to fantasized violence also proved to be high and low in overall exposure to television. Of the 20 programs rated by Greenberg and Gordon's panel as least violent, the median number viewed very frequently by the high and low groups was, respectively, 9.5 (range 3-17) and one (range 0-11). The difference is significant ($p < .01$, two-tail).[4] Unlike the viewing of violent television programs, however, the viewing of nonviolent programs does reveal some overlap between the two groups; five of 20 low-exposure subjects viewed within the range of the high-exposure group. Another questionnaire item indicated that the low- and high-exposure subgroups differed in more than sheer exposure to televised fantasy of any sort: they also differed in a preference for violent fantasy. Subjects were asked to indicate their three favorite television programs. 60 choices were made by high-exposure subjects, 59 by low-exposure subjects. For high-exposure subjects, 27 of these choices were among Greenberg and Gordon's 20 most violent programs; for low-exposure subjects, only eight choices were programs in the same group. The programs receiving the most votes among high-exposure subjects were *Mission Impossible* (eight votes) and *Hawaii Five-O* (five votes); the programs receiving the most votes among low-exposure subjects were *World of Disney* (five votes) and *Brady Bunch* (four votes).[5] Of 65 shows rated by the Greenberg and Gordon panel of citizens with mean ratings of violence ranging from 3.56 to 1.06, *Mission Impossible* was ranked third (3.35), *Hawaii Five-O* fourth (3.24), *World of Disney* twenty-seventh (1.56), and *Brady Bunch* fifty-first (1.17). For the 20 least violent shows, there were six votes from high-exposure subjects and 12 votes from low-exposure subjects.

Another way of classifying subjects' favorite programs is presented in Table 1. It is clear that the low-exposure subjects show a relatively greater liking for situation comedies, musical-variety programs, and nature series than do high-exposure subjects. In the same table, data also are tabulated from a question asking subjects how many and what comic books they regularly read. The comic book preferences seem to be in agreement with the favorite-program data in indicating a greater interest in violent material among the high-exposure subgroup. Thus several lines of evidence converge in suggesting that the two subject groups differ *both* in overall television viewing *and* in preference for specifically violent media fantasy.

Table 1: Media preferences of subjects high and low in exposure to televised violence

A. Three favorite television programs: number of votes received		
	High exposure subjects	Low exposure subjects
Situation comedy	4	11
Comedy-variety	7	5
Family series[1]	4	5
Musical-variety	0	6
Sports	0	1
Adult series[2]	10	7
Violent drama	26	9
"Humorous-violent"[3]	3	3
Children's-nature[4]	2	8
Afternoon children's series	4	4
B. Comic books: number of subjects reporting		
Regularly read at least one comic	13	12
Regularly read at least one humorous comic	7	9
Regularly read at least one violent-hero comic	7	4
Report reading *only* violent-hero comics	6	2

[1] I.e., *Brady Bunch, Courtship of Eddie's Father*

[2] Either serious or humorous, e.g., *Marcus Welby, Love American Style*

[3] I.e., *Get Smart, Hogan's Heroes*

[4] I.e., *World of Disney* (variable in content from week to week but classified here), *Wild Kingdom*

At the time subjects took the screening test, an interviewer ascertained from each child or the accompanying parent where the head of household worked and what he or she did there. On the basis of this information, each child was assigned a social class score based on Warner et al.'s Revised Scale for Rating Occupations (Warner, Meeker, and Eells, 1960). The median value for the high-exposure group was 3; for the low-exposure group it was 2 (.10 > p > .05, two-tail). These values roughly correspond with lower- and upper-middle class occupations.

Working- or lower-class occupations (ratings of 5-7) were more frequent among the high-exposure group (seven children *vs* two in the low-exposure group), while middle- or upper-class occupations (ratings of 1-3) were slightly more frequent among the low-exposure group (15 children *vs* 12 in the high-exposure group). Thus both groups were predominantly middle class, but with a greater representation of the working class in the high-exposure group and of the top professions and entrepreneurial positions in the low-exposure group. (Five low-exposure subjects and two high-exposure subjects received a rating of 1.) The overlap of the two groups (a range from 1 to 7 within each), however, suggests only a very imperfect correlation of social class even with these extreme groups in terms of degree of exposure to televised violence.

No measures of the subjects' intelligence were available. The screening questionnaire did, however, ask subjects to indicate their typical grades in school. The question gave eight scaled alternatives ranging from "all As" (assigned a rank of 7) to "sometimes Fs" (assigned a rank of 0). The two groups did not differ significantly in responses to the question (a median ranking of 4.5 for the high-exposure group vs 5 for the low-exposure group where 4 = "Bs and Cs" and 5 = "mostly Bs"; $p > .10$, two-tail).

A series of questions on the screening test gave six scaled temporal alternatives ("no time at all" to "four or more hours a day") for amount of interaction with various categorizations of family, peers, and siblings (parents together, mother only, father only, group of friends, single good friend, alone, with sibs).[6] Most subjects in both groups reported spending: more time with a group of friends than alone (12 *vs* six low exposure, 13 *vs* five high exposure); more time with a single friend than alone (13 *vs* six low exposure, 14 *vs* six high exposure); more time with parents together than with a group of friends (11 *vs* five low exposure, 14 *vs* four high exposure) and more time with parents singly or together than with friends singly or together (11 *vs* seven low exposure, 14 *vs* five high exposure). The groups did not differ significantly in absolute time reportedly spent with peers, either in groups or singly, or in absolute time spent alone or with sibs. The low-exposure subjects did report significantly less time spent with father alone ($p < .05$, two-tail) or with the parents together ($p < .02$, two-tail) than did the high-exposure subjects, but did not differ in time reportedly spent with mother. In relative terms, four low-exposure but nine high-exposure subjects reported spending more time with fathers than with mothers. The peer data give no indication of differences in social adjustment for the two groups, and the father-contact difference may merely reflect the social class composition of the two groups if it can be assumed that higher-status professionals and entrepreneurs have relatively less time to spend at home. It is interesting to note, however, that the group which prefers violent fantasy also seems to have relatively more exposure to the father, the parent whom cultural stereotypes indicate should be more "aggressive." It must be recognized that all these data measure the quantity, rather than the quality, of social relations.

The median number of sibs for high-exposure subjects was four; for low-exposure subjects, it was three. This difference approached marginal significance ($p \cong .10$, two-tail). Birth order did not seem to vary: each group contained six first-borns, with two last-borns among the low-exposure subjects and three last-borns in the high-exposure group.

Several questions about reported "disturbances" of sleep were used as possible indicants of psychological distrubance. An index of "minor" sleep disturbance (giving high weights to reports of relatively long sleep latencies and of relatively many spontaneous awakenings in home sleep) did not discriminate the two groups. Neither did a scale relating to the reported frequency of home nightmares. Other "major" disturbances of sleep (enuresis, somnambulism), however, were appreciably but not significantly more often reported by high- than by low-exposure subjects. Three high-exposure subjects reported two to three enuretic episodes per week. The most serious enuretic problem in the low-exposure group was one subject who reported enuresis "once a month or so." Seven subjects reported a history of somnambulistic episodes; six were from the high-exposure group.

It might be wondered whether the high and low groups differed in exposure to televised violence as a matter of personal choice or in response to differential parental restrictiveness. We have no directly pertinent data. However, if one imagines that parental demands are likely to vary systematically in strength and effectiveness during the course of the week, there is a possible indirect test for level of parental demands. It seems reasonable to assume that parental demands on schoolchildren's viewing are likely to be somewhat relaxed on Saturdays and to be less obligatory and well-policed during weekday afternoons than during the evening hours, when both parents are home and when cultural norms indicate that the child should be doing his or her homework. Thus if the low-exposure subjects are under greater parental restriction than high-exposure children, one might predict that: (1) low-exposure subjects will watch relatively more of their own quota of frequently viewed programs on weekday afternoons and Saturdays than will high-exposure subjects and consequently, (2) there should be less difference in viewing between the two groups at these times than during the remainder of the week. These predictions are borne out by the data. The low-exposure group watches 40 percent of its own quota of frequently viewed programs on weekday afternoons and Saturdays, compared to a 26 percent figure for the high-exposure group. The between-groups ratio of frequently viewed programs drops from approximately 4.5:1 during weeknights (means of 28.65 *vs* 6.10 programs frequently watched) to approximately 2.5:1 on weekday afternoons and Saturdays (10.15 *vs* 4.05 programs frequently watched). These data are susceptible to other interpretations, of course, but they at least do not seem inconsistent with the hypothesis of greater parental regulation of television behavior in the low-exposure group.

Two questions on leisure activities allowed some assessment of what might be called an "aggressive" attitude toward natural phenomena and, in addition, sought to isolate whether this attitude was purely personal or received active familial support. Subjects were asked, on three-point scales (0 = never; 1 = sometimes; 2 = often) to indicate how regularly they hunted and fished alone and with their families. Both hunting and fishing are significant recreational activities in the small Rocky Mountain community in which the present study was conducted: only one of 40 subjects claimed that he never fished, while 26 subjects said that they hunted at least occasionally. High-exposure subjects reported engaging more often in hunting and fishing than did low-exposure subjects (mean scale differences of 1.05 vs .65 and 1.70 vs 1.25, respectively). The difference for fishing was significant ($p < .02$, two-tail), while that for hunting only approached a marginally significant value ($p \cong .10$, two-tail).

An analysis was conducted of subjects with differential levels of hunting and fishing when alone *vs* when with their families. For hunting, 50 percent of 12 high-exposure subjects reported the activity more often with their family than alone, and 50 percent reported it more often alone. Comparable figures for eight low-exposure subjects were 25 percent with family and 75 percent alone. For fishing, however, 64 percent of 14 high-exposure subjects reported more of the activity with the family, while 78 percent of nine low-exposure subjects reported more in the family context. While these differences are not of appreciable magnitude, they suggest that low-exposure subjects received little familial support for hunting and more for fishing, while high-exposure subjects received about the same level of support for both activities. If fishing is considered potentially more contemplative and less hostile than hunting (Gottschalk, Winget, and Gleser, 1969), then this in turn suggests differences in "aggressive style" between the families (presumably between fathers, in particular) of subjects in the two groups which are consistent with their sons' viewing preferences.

In summary, compared with the low-exposure group, boys in the high-exposure group tended to: be higher in overall television consumption; show greater enjoyment of violent media content; be of somewhat lower social class background; come from somewhat larger families; report more frequent contact with their fathers; report a slightly higher incidence of enuresis and somnambulism at home; possibly experience less parental supervision of television behavior; hunt and fish more frequently and come from families that may be relatively more sympathetic to hunting. Boys in the two groups did not seem to differ in self-reported academic achievement, in peer group interaction, in birth order status, or in the incidence of frightening dreams at home. Some of the positive findings are consistent in direction with those of earlier research, but they seldom achieved sufficient magnitude to be very useful in a predictive sense. In the present context, they are best viewed simply as descriptions of the current samples rather than as generalizations to other children with comparable differences in degree of habitual exposure to televised violence.

Orientation. Subjects were told that the purpose of the study was to make comparisons of sleep and dreams recorded in the University's old Sleep Laboratory facility, housed in the building in which they took their screening test, and in its new Sleep Laboratory facility in a building whose construction had just been completed. They were informed that the experimenters were interested in what effect the new setting might have on children involved in an ongoing longitudinal study (Foulkes, 1969) who soon would make the shift from the old to the new facility. (That we had been studying children's sleep and dreams for some time was well known to most of the children and had been stressed in our newspaper advertisement.) As a rationale for the study in which they were to serve, subjects were told that we wanted to test the setting factor first with them before committing our long-term subjects to the changed laboratory environment. The screening test was explained to the subjects in terms of the goal of using them to predict effects for the longitudinal group: they were to be selected on the basis of how well their answers matched those of boys already under investigation in the longitudinal studies.

In line with this rationale, all subjects spent an initial night in the old Sleep Laboratory for adaptation to electrophysiological recording during sleep, to nocturnal dream interviews, and to other constant features of the laboratory regime. Adaptation nights were run three to four weeks before the experimental nights. The experimental treatments (films and involvement manipulations) were administered on two nights, seven to nine days apart, in the new Sleep Laboratory.[7]

The change in setting from adaptation to experimental nights seemed to contribute substantially to the credibility of the stated rationale for the study and also added to the credibility of manipulations performed in the new laboratory. The television sets for subjects' viewing of films, for example, rather than seeming to be specifically related to the current study, were interpretable simply as an integral part of a newly equipped and more modern laboratory facility. Likewise, a closed-circuit television system (cameras in the subjects' rooms and monitors outside), which was installed after the first 21 subjects had completed their service, was explained to later subjects as another integral feature of the new laboratory facility, enabling experimenters to observe the subjects' body positions during sleep without having to enter the room. Most subjects seemed favorably impressed by the new laboratory and agreed it was a marked improvement upon the older makeshift facility. Subjects did not voice suspicions of the rationale presented to them either for the study as a whole or for particular manipulations within it.

This situation contrasts with that in the prior study of preadolescents by Foulkes et al. (1967), in which there was generally lower manipulation credibility (subjects were told films would be shown to them to keep them occupied while the polygraph "warmed up"; several voiced suspicions of the real reason for the film administration) and a more obtrusive film manipulation (a projectionist showed individual subjects the films in a small room adjacent to the sleeping area). In Foulkes and Rechtschaffen's (1964) study of young adults, films also were shown individually by a projectionist, and no effort was made to disguise the purpose of the film administration.

Laboratory routine

Experimental procedures. Subjects were assigned standard times for coming to the laboratory. Upon arrival there, they changed into their bedclothes. They then were taken to a control area, where an experimenter attached electrodes to their faces and scalps for EEG (electroencephalogram) and EOG (electro-oculogram) recordings. The same two experimenters were present during the presleep period on all nights, and the same experimenter was assigned the task of attaching electrodes and otherwise attending to a given subject on all three nights of his service in the study. Boys were run together in the same groups of three on each laboratory night, so that presleep peer companionship also was a common factor across nights.[8]

Electrophysiological recordings were taken each night until the completion of subjects' schedule of REM-sleep awakenings for dream retrieval. Recording began at lights-out on adaptation nights; on experimental nights, it began at the time at which the film administration was scheduled. Four channels of information were recorded polygraphically for each subject: two channels for horizontal EOG tracings, one for a vertical EOG and prefrontal (F_p2) EEG tracing, and one for a central (C4) EEG tracing. Sleep variables (latency from lights-out to sleep onset, sleep latency to REM-sleep onset, number of spontaneous awakenings of greater than three minutes' duration, and cumulative time spent in such awakenings) were scored by one experimenter during the course of each night. For ease of on-the-spot scoring, the onset of REM sleep was identified by the first REM, rather than by the onset of its EEG manifestations.

Awakenings for dream retrieval were scheduled for ten minutes after the onset of the first three REM-sleep periods achieving that duration on the adaptation night and ten minutes after the onset of the first four REM periods of such duration on experimental nights. On adaptation nights, one of the experimenters conducted the interviews. On experimental nights, awakenings were initiated, on a signal from the experimenter, by an interviewer who had arrived at the laboratory after lights-out and was ignorant of each subject's status on exposure, film, and involvement variables. A given subject was awakened by the same interviewer on each of his experimental nights.[9]

Interviewers followed a standard format. After spontaneously reporting his dream, the subject was questioned about: the presence and nature of the visual imagery; the hallucinatory quality of his experience; dream characters and their age, sex, and familiarity to him; his own role in the dream experience; dream settings and their familiarity to him; and feelings associated with the dream experience. In the case of thoughtlike experiences (Foulkes, 1962), some of these questions were omitted as they did not apply. If the subject initially could not recall anything, he was asked to "think for just a moment" (30 seconds) to see if he could remember anything. If not, he was asked whether he had had a dream but forgotten it or whether it seemed like he simply hadn't been dreaming before his awakening. If he recalled something after the delay, the regular format was followed. The interview questions were designed to clarify dream characteristics of general significance, and also to collect information specifically pertinent to the application of dream scoring devices later to be employed by judges. All interviews were conducted over an intercom system independently connecting an audio control room to each of the three bedrooms, and were recorded for later transcription.

Experimental nights were routinely terminated at about 7 a.m. whether or not a subject had completed his awakening schedule. On eight of the 80 experimental subject-nights, only three awakenings could be made. These nights were distributed 5-3 between incidental-focal and between nonviolent-violent conditions; they were distributed 6-2 between low- and high-exposure subjects. Failure to complete the schedule might create a presumption of disturbed sleep, but these data indicate no marked effects of the independent variables on laboratory sleep patterns. This conclusion is supported by more formal analyses of the four EEG sleep variables which were scored. There was only one significant main effect out of the 12 possible: more time awake following spontaneous awakenings in incidental than in focal involvement ($\bar{X}$ = 13.38 min. *vs* $\bar{X}$ = 11.08 min., F = 4.668, 1/34 df, $p < .05$). Likewise, only one of 12 first-order interactions was significant—that for number of spontaneous awakenings and between involvement and prior exposure rate (F = 4.384, 1/34 df, $p < .05$). More such awakenings were scored for low-exposure subjects under incidental involvement ($\bar{X}$ = 2.40) and for high-exposure subjects under focal involvement ($\bar{X}$ = 1.10). One triple interaction, that for REM latency, was significant (F = 9.454, 1/34 df, $p < .005$). High-exposure subjects had their briefest latencies and low-exposure subjects their longest latencies under violent-focal and nonviolent-incidental conditions, while violent-incidental and nonviolent-focal conditions were associated with relatively long latencies for the first group and relatively short latencies for the second group.

Films. Two episodes, one relatively violent and the other relatively nonviolent, were obtained from a "children's western" (Schramm et al., 1961) series.[10] Although this series had not been included in the listing of current television series on the screening questionnaire, it had been shown regularly on a Denver network affiliate during the preceding year. The series was rated among the 20 most violent television series by Greenberg and Gordon's (1970) sample.

The two episodes originally had been screened during the season immediately preceding the summer in which the present study was conducted. In apparently informal conversation with subjects on the mornings following experimental nights, an experimenter ascertained that 14 subjects recalled having seen, or thought they had seen, the nonviolent episode before, while 13 subjects recalled having seen, or thought they had seen, the violent episode before; 16 of these recalls were by high-exposure subjects (nine violent, seven nonviolent), and 11 were by low-exposure subjects (four violent, seven nonviolent). No subject, however, evidenced any disappointment or voiced any complaints about seeing a particular episode over again, and their "recall" of having seen the films before seemed vague at best.

The fact that the episodes came from the 1969-70 season and from a "children's western" series correctly indicates that aggressive displays in the violent film were relatively "sanitized" compared, for example, to those in an earlier film from an adult western series used as a violent stimulus by Foulkes and Rechtschaffen (1964). In relative terms, however, the two stimuli used in the present study were markedly divergent in their reliance on violence and threat. A tabulation was made for each film of the number of its violent "elements" (continuous scenes of malicious physical violence to persons or verbal or gestural threats thereof directed by the same party or parties to the same party or parties). The nonviolent stimulus was judged to contain three such elements, the violent film 26.

The nonviolent stimulus was a light comedy about an imposter deliberately sent to a frontier town in place of an invited but travel-weary French prince. Much of the play dealt with the excitement produced by this visit from "royalty" and with the schemes of two mothers to make a favorable match for their daughters with the supposed prince. At the end, the real prince arrives and finds himself overwhelmed by the townspeople. The likeable imposter is forgiven, and makes the match of his choice.

The violent stimulus portrayed a greedy and eccentric old man who plans to rob a valuable cargo of furs. A boy catches wind of this and races to alert his father, who is part of the convoy escorting the furs.

The son is captured, and when his life is threatened the father surrenders the cargo (at which time a box carrying gold bars is discovered beneath the furs). The remainder of the film showed the pursuit and capture of the old man and his companions. Significant episodes included one in which the boy manages temporarily to gain the upper hand and threatens the lives of the robbers and another in which the boy is physically mistreated for this display of bravery. At the end, "justice triumphs." An adult male hero and a juvenile male hero were present in both the nonviolent and the violent films, although they were less central to the plot of the former. Both films were shown complete with commercial messages but were terminated at the onset of a preview of the next episode.

Involvement. On all nights, subjects were told that the study required that physiological recordings begin at 9 p.m. On experimental nights, subjects were also told that they would be asked to stay awake during the first hour of recording so experimenters could obtain samples of their waking brainwave patterns to compare with their sleeping patterns. In both the focal and the incidental involvement conditions, the experimenters indicated some concern that the subject have some activity to engage his interest during this presleep recording period. The manipulation consisted of differential emphasis upon television viewing as such an activity.

In each of the three bedrooms, a television receiver was mounted to the wall at the foot of the subject's bed. Subjects generally noted the sets upon their arrival at the laboratory and asked if they could watch television. They were told that although the new laboratory facility was not yet connected to an outside cable, the receivers were connected to the University's television station, which broadcast at irregular times. The experimenters said they understood that a program would be on later during the evening, although one could never be sure just what it would be.

As the subjects were put to bed after the attachment of their electrodes, the experimenter turned on the receiver, which generally showed a test pattern along with a background of recorded music. Subjects were told that a program probably would come on within a few minutes. Arrangements had previously been made with the University television station to have our films shown each evening at about 9 p.m., after a call from the Sleep Laboratory to the television studio.[11] As soon as all subjects were abed, the call was made and the television program promptly began.[12]

All subjects on a given night saw the same film. The order in which experimental nights were run was such that all variables—prior exposure rate, film order, involvement—were randomized over the time-course of data collection. Experimental nights were run in two "series," with 21 subjects in service between August 3 and 17 (with the exception of the two aforementioned "delinquents," who completed the series on August 22) and 21 subjects in service between August 18 and September 2.

The only major difference between the two series was the absence or presence of the closed-circuit, subject monitoring television system. Our original intention was to monitor all subjects over the system during the presleep recording period in order to validate the involvement manipulation. When the equipment for this system did not arrive by the scheduled onset of the experimental nights, it was decided to begin with another form of behavioral observation. Both our subjects and experimental assistants faced later school obligations that would have been in serious conflict with the performance of their experimental chores had the experiment been delayed beyond this point.

The *focal involvement* manipulation was accomplished as follows: Subjects watched the films while seated in bed with low or no other room illumination, with their door closed, and with the television receiver directly facing them and tuned in at moderately high volume. No other materials were available in the room to distract them from television viewing. They were told: "Tonight we'd like to have you stay awake and just relax for the first hour of recording, and then we'll let you go to sleep. Why don't you just watch the television for this time—there'll be a show on soon. We'll come in and turn the set off when it's over and we're ready to have you try to go to sleep." The subject's door then was closed, and there was no further interaction with the experimenter until the program was over. If a subject did call out, over an intercom, that he wanted the experimenter, attempts were made to handle whatever was on the subject's mind without entering his room. Continuous physiological monitoring established that the subject remained awake, and the intercom monitoring established that the sound portion of the television program was still on at the desired volume during the presleep period. Subjects were told to notify us of any picture difficulties, and the experimenter, in conversation with the subject at the end of the presleep period, verified that the program had been adequately received in all cases.

The *incidental involvement* manipulation was implemented as follows: Subjects watched the films while seated in bed with high room illumination, with the door from their room to the control area open; thus they were aware of the noise of the polygraph recorders and of the experimenters moving about outside. The television set was a slight angle away from a direct line of gaze, and the volume setting on the receiver was at a relatively low but audible volume. Before leaving the subject's room at the start of the presleep period, the experimenter carried in a jigsaw puzzle map of the United States, several recent issues of *Boys' Life* and of science-mechanics magazines, an illustrated nature book, and a Soma Cube Puzzle.[13] He told the subject: "Tonight we'd like to have you stay awake and pretty alert for the first hour of recording, and then we'll let you go to sleep. Now, what you do in the hour is up to you. There's the TV there [motions in direction of set] you may watch if you want to, and here are some magazines and a [map] puzzle and book that's got some real nice pictures in it [shows some], and then here's a puzzle [Soma] that's kind of fun if you want to try it [demonstrates]." The experimenter then left. At three prearranged points in relation to the plot of the films (at moments of high excitement during the violent film and at rather unexceptional moments of the nonviolent film), the experimenter entered the room, ostensibly to see how the subject was getting along ("Are you managing to stay awake OK?") or to inspect the electrode placements. Brief, informal conversation occurred at these points. If, as sometimes happened, a subject called out to the experimenter, the latter entered the room to see what the subject wanted. Many of these unscheduled entries were occasioned by a subject's success or lack thereof in working the Soma problem, which was the most successful device for distracting a subject's attention from the television program.

Validation of involvement manipulations. The validation of the involvement manipulation had to proceed differently during the first series of experimental subject-nights (n = 21), in which closed-circuit subject monitoring was not available, and during the second series of experimental subject-nights (n = 21), in which it was available. One subject in each series was excluded from the final data analysis, for reasons already noted. In addition, one subject in the second series could not be observed via television due to equipment failure, so that, of the effective subject sample of 40, 19 were observed via television monitors.

The other 21 subjects were studied with less refined behavioral observation techniques. Focal-condition (door-closed) subjects could not be observed without distracting them in a way detrimental to the manipulation. This did not seem of great concern, however, since there was little else for focal-condition subjects to do but watch television (or fall asleep, which we did have the means to monitor physiologically and which, in fact, never did occur). The incidental manipulation, where subjects had a choice of activities other than television, was the more critical situation for behavioral monitoring. Fortunately, any behavioral observations contributing to distraction from ongoing activities were compatible with it. Also, the incidental manipulation offered several opportunities for experimenters to observe the subjects, both on their prearranged trips into subjects' rooms and on random trips in the control area which carried them past the open doorways to these rooms.

For these 21 subjects, the experimenter assigned an involvement rating for attentiveness to television on a six-point scale, as follows:

6—Seems to have watched television almost continuously
5—Watching most of the time; minor attention to other activities
4—Watching more than not; some attention to other activities
3—More attention to other activities; some watching of television

2—Most attention to other activities; minor attention to television
1—Seems hardly to have watched at all

Scores assigned to experimental subjects on 21 incidental-viewing nights observed in this fashion were as follows:

Rating	*Number of Nights*
6	2
5	0
4	3
3	6
2	4
1	6

The distribution reflects the inherent variability of a condition in which subjects had a meaningful choice about whether or not to watch television, but it also indicates that the incidental manipulation generally was successful in producing less than maximal attention to the program. (Sixteen of 21 subject-nights were associated with ratings of more attention to other activities than to television). Of course, the use of these ratings to validate the incidental-focal manipulation cannot be conclusive since: (1) focal ratings are lacking, and it is only presumed that viewing is more intent in that condition; (2) the ratings for the incidental condition are by the same experimenters performing the manipulations creating the condition.

For this reason, it is important to note that results consistent with the findings and presumptions of the behavior ratings of the first series were obtained from the second subject-series in which both focal and incidental nights were observed via television monitors by three other experimental assistants.[14] On these nights, the experimenters closed the doors to the rooms of all subjects after they introduced the focal or incidental manipulation. A signal then was given to the observers, who quietly entered the control area and positioned themselves before the monitors. A subject's door was reopened as the film began if he was in the incidental involvement condition. The observers left after the experimenters closed the doors as they entered subjects' rooms to turn off the television receiver. There was no indication that any subject was aware of the observers' presence.

The observers watched the television monitors continuously during the film presentation. The observation period began with a signal from an experimenter coordinated with the onset of the lead-in to the first act of the film, and ended with a similar signal coordinated with the onset of the commercial message following the epilogue of the film episode. The observers were not systematically informed of the reason for the open vs closed door condition, nor of other variables differentially associated with the involvement manipulation, although they undoubtedly formed their own ideas about the different conditions in which subjects were observed. Observers also were not informed of the details of the stimuli being used, although they could, at least in the open door condition, hear some of the film dialog and action. One observer knew the exposure-variable status of the subjects; the other two did not. Each observer watched the same subjects on both experimental nights.

The observation schedule called for an assignment of every 15-second period to one of the following categories relating to the subject's eye contact with the television set: continuous, 10-15 seconds, 5-10 seconds, 0-5 seconds, none. A space was also provided on the observation sheets for comments about the subject's body attitude and motility and about his activity if he was not watching the television receiver. A large wall clock with a sweep-second hand or a stopwatch was used for marking off the 15-second intervals.

The validity of the involvement manipulation was tested by finding, for each subject-night, the percentage of 15-second observation periods judged to contain either continuous or 10-15 seconds ("nearly continuous") eye contact. Averages of individual subject-night percentages then were computed for the focal and incidental involvement conditions. On 19 focal subject-nights, the average of continuous or nearly continuous eye contact with the television receiver was 97.5 percent (range of individual subject-night values: 89 to 100 percent); on 19 incidental subject-nights, the corresponding average was 21.2 percent (range of individual subject-night values: 1 to 63 percent). Thus the focal and incidental conditions were markedly different in the absolute quality of viewing behavior elicited, and there was no overlap whatsoever between viewing behavior in the two conditions. Seven subjects were observed in both conditions: their median focal vs incidental percentage *difference* in continuous or nearly continuous viewing was 76 percent, with a range from 37 to 94 percent. It can be concluded that the experimental manipulation of viewing involvement was highly successful.

Involvement interaction with other variables. It might be asked, however, whether the involvement manipulation was equally successful for high- and low-exposure subjects or for the violent and nonviolent films. It would be important to know, for example, whether any interactions which subsequently might be observed between the involvement manipulation and subject or film variables could be attributed to the different degrees to which the manipulation worked for different subjects or films or whether such effects occurred in spite of equivalent amounts of actual viewing behavior for different subjects or films. In the former case, the effects could be attributed to quantitative variability of viewing behavior within a given involvement condition; in the latter case, the effects presumably would have to be attributed to the quality of viewing behavior. It also seems of general methodological interest to know whether an act-of-viewing variable can be manipulated in equivalent fashion for different kinds of media stimuli and for subjects with different habitual levels of interest in typical media stimuli.

Data from the behavioral ratings of the first series of experimental nights showed no significant or appreciable difference between low-exposure subjects (mean rating for ten subject-nights: 2.70) and high-exposure subjects (mean rating for 11 subject-nights: 2.64) in the level of film viewing involvement created in the incidental condition. The difference between violent film, incidental-viewing nights ($N = 11$, $\overline{X} = 3.00$) and nonviolent film, incidental-viewing nights ($N = 10$, $\overline{X} = 2.30$) was somewhat larger but not significant. Data from the television-monitored scoring of the second experimental series indicated no significant or appreciable differences in the mean subject-night percentages of continuous or nearly continuous viewing between high-exposure subjects (Focal: 96.6 percent, range 89-100 percent; Incidental: 21.2 percent, range 6-63 percent) and low-exposure subjects (Focal: 98.4 percent, range 93-100 percent; Incidental: 21.1 percent, range 1-54 percent) or between the violent film (Focal: 97.1 percent, range 89-100; Incidental: 16.2 percent, range 1-44 percent) and the nonviolent film (Focal 98.0 percent, range 95-100 percent; Incidental: 25.6 percent, range 6-63 percent). Overall, then, it can be concluded that the involvement manipulation "took" equally well for both subject groups and for both films.

Analysis of dream data

Magnetic tape recordings of all nocturnal interviews were transcribed by a typist, who then removed all identifying information from each interview transcript and randomly assigned it a number from 1 to 320. Judges received and scored the interviews in this numerical order. Generally both judges for a dream variable were naïve with respect to the association of any given transcript with subject or treatment variables. In these cases, two-rater averages were employed in subsequent data analysis. For certain scores, noted below, one judge was naïve in this sense and one was not. Here the score of the naïve judge was employed in subsequent data analysis, the other score used only to establish the reliability of the first judge.[15]

The dimensions along which dreams were judged fit into four general areas: general dream intensity or vividness; hedonic quality of the dream experience; aggression and hostility; and defensiveness or guilt over impulse expression. The first three areas were suggested by prior research indicating that violent films may increase the overall intensity of dream experience without affecting hedonic tone or the expression of aggression (Foulkes and Rechtschaffen, 1964) or may reduce dream intensity, "bad" dreams, and aggressive dream content (Foulkes et al., 1967). The fourth area was included to determine if any reductions in aggressive material that might be observed in dream content could be attributed to increased vigilance of defenses against aggression, as manifested in increased dream distortion or in manifest anxiety. The specific measures employed in each of these four areas are discussed briefly below, along with the interjudge reliabilities.

General dream intensity. 1. Dreamlike Fantasy Scale. This eight-point scale "originally was constructed on the basis of observations of mental activity at sleep onset indicating a progression from 'thought-like' mentation to pictorial mentation to hallucinatory dreamlike imagery" (Foulkes, 1970). It has been used with some success in a number of studies as an index of overall dream intensity (Hauri,Sawyer,and Rechtschaffen, 1967). Another presumed component of the same factor is "imagination," which Foulkes et al. (1967) found to be reduced in level following the viewing of a violent film. The Dreamlike Fantasy (Df) scale was applied here by two naïve judges. Pearson r for the interjudge reliability of individual report classification was .97. This scale is applied to all awakening protocols and in this sense parallels the Goodness of Recall scale found by Foulkes et al. to discriminate dreams obtained following violent and nonviolent films; all other scales mentioned below apply only to interviews which produced some substantive dream content. The report N in the first case is 312; in the latter case, it is 227.

2. Dream Word Count. Hauri et al. (1967) found that dream word count was highly loaded on their "vivid fantasy," or general dream intensity, factor. This variable was scored here by only one judge, following explicit rules developed expressly for children's laboratory dreams (Foulkes and Shepherd, 1971). The judge was not naïve but had indicated in the prior study that she could achieve high interjudge reliability applying these same rules.

Hedonic quality. 1. Hedonic Tone Scale. This seven-point scale is similar to that of Foulkes and Rechtschaffen (1964); it has been used by Foulkes et al. (1969) and by Weisz and Foulkes (1970). Its values extend from 1 (= very pleasant) through 4 (= neither) to 7 (= very unpleasant). Hedonic tone is defined at the level of the entire dream protocol. The scale was applied here by two naïve judges. Pearson r for interjudge reliability of individual report classification was .81.

2. Unfavorable Outcome. This classification comes from the Foulkes-Shepherd (1971) scoring system for children's dreams. It refers to experiences of victimization, failure, and bad fortune, and is scored for each dream on a present *vs* absent basis. It is scored for (a) dreamer, (b) others, and (c) either (dreamer and/or others). It was scored here by one naïve and one nonnaïve judge. The reliability estimate we have employed is the percentage of positive scorings of the attribute by both judges which were concordant (Yarrow, Campbell, and Burton, 1970). In the present scoring, the reliabilities were 83 percent (dreamer), 93 percent (others), and 91 percent (either).

3. Favorable Outcome. This classification also comes from the Foulkes-Shepherd (1971) system and refers to experiences of being the recipient of friendly acts from others, of success, or of good fortune. It

too is scored on a present-absent basis for (a) dreamer, (b) others, and (c) either. Reliabilities for the same two judges as above were 80 percent (dreamer), 84 percent (others), and 85 percent (either).

Aggression and hostility. 1. Physical Aggression Scale. This seven-point (0-6) scale was first used by Foulkes and Rechtschaffen (1964) in their film-dream study. It has since been found useful in studies by Foulkes et al. (1969) and Weisz and Foulkes (1970). In the present study it was applied by two naïve judges. Pearson r for interjudge reliability of individual report classification was .85. (Originally we planned to use the Foulkes-Rechtschaffen Verbal Aggression Scale as well, but aggression so rarely took a verbal form in the boys' dreams—only three dream reports of 227 being judged > 0 by both judges—that this was not feasible.)

2. Attack Social. This category, scored on a present-absent basis for each dream, is taken from the Foulkes-Shepherd (1971) scoring system. It comprehends both verbal and physical negative actions and is a refinement of a scoring category (Social: Escape/Hostile) found by Foulkes et al. (1967) to be diminished in children's dreams following the viewing of a violent film. The category is scored for (a) dreamer, (b) others, and (c) either. It was scored here by one naïve and one nonnaïve judge. Reliabilities, computed as for other variables from the same system, were 95 percent (dreamer), 88 percent (others), and 93 percent (either).

3. Approach Social. This category comes from the Foulkes-Shepherd (1971) scoring system and also is a refinement of a content analysis category in the earlier film-dream study of Foulkes et al. (1967). It refers to friendly, supporting, or playful overtures to others. It too is scored on a present-absent basis for each dream, and is scored for (a) dreamer, (b) others, and (c) either. Reliabilities, computed as above and for the same two judges, were 87 percent (dreamer), 87 percent (others), and 90 percent (either).

4. Hostility Need. This five-point (0-4) scale was taken directly from the prior film-dream study of Foulkes et al. It refers to hostile feelings or actions initiated by the dreamer himself. It was scored by two naïve judges, with a Pearson r for reliability of individual report classification of .83.

5. Hostility Press. This five-point (0-4) scale was also taken directly from the study by Foulkes et al., in which it significantly differentiated the violent and nonviolent presleep film conditions. It refers to hostile feelings or actions initiated by others (not necessarily toward the dreamer). It was scored by two naïve judges, with a Pearson r for reliability of individual report classification of .83.

Defensiveness-guilt. 1. The Anxiety Scale. This scale was developed and has been used extensively by Gottschalk and associates (Gottschalk and Gleser, 1969; Gottschalk et al., 1969). The scale includes six subscales, yielding a total score on the assumption that different subclasses of anxiety are of equal importance in determining a person's overall anxiety level. Scoring includes the application of a correction procedure for differences in overall protocol length. We have employed the total scale score and that of those subscales which seemed conceptually most relevant to aggression anxiety: (a) total scale score; (b) death anxiety; (c) mutilation anxiety; (d) guilt anxiety.

In view of the authors' suggestion that a 70-word minimum sample is required for the application of their scale, the entire dream protocol of a subject-night was the unit of analysis here. (For other variables, we treated individual dream reports as a unit, and averaged scores for such reports to obtain an overall score for a subject-night). Scoring here was by two naïve judges. The reliabilities (Pearson r) of their corrected scores for subject-night protocols were .88 (total scale score); .79 (death); .58 (mutilation); .62 (guilt).

2. Representativeness of Setting. This and the following scale are taken from the Foulkes-Shepherd (1971) scoring system and are measures of the degree of distortion of major dream elements. Representativeness of Setting is scored as 1 if the setting is known and familiar, 2 if it is known but changed, and 3 if it is unknown or vague. The scale is applied to individual dream reports. It was scored here by two naïve judges. Reliability, as determined by the number of dreams on which exact interjudge agreement was achieved, was 92 percent.

3. Representativeness of Characterization. This is a four-point scale (1 = all familiar; 2 = some strange; 3 = all strange; 4 = no characters). Scores of 4 are not considered continuous with those of 1-3 and are discarded in data analysis ("familiarity of characters" does not apply when none are present), but they are used in determining reliability of interjudge classification. This scale also is applied to individual dream reports. Scored by the same two judges as above, the percentage of exact agreement was 92 percent. For two dreams judged as 4 by one judge but as scoreable by the other judge, we examined the protocols, determined that they were in fact scoreable, and resubmitted them to the first judge for scoring. His incorrect judgments of 4 were used in computing reliability, but his rescorings were used in data analysis.

4. Guilt. This six-point (0-5) scale, taken directly from Foulkes et al. (1967), already has been discussed extensively. It was scored here by two naïve judges, with a Pearson r for individual report classification of .88.

RESULTS

Recall. Overall, dreams were recalled on 227 of 312 experimental night awakenings from REM sleep, for a pooled recall percentage of 73 percent (This is highly similar to the 69 percent reported by Foulkes et al. [1967] for their 10-12-year-old boys.) At least one dream was reported on each of the 80 subject-nights. This degree of dream recall is more than sufficient to allow the testing of hypotheses relating to dream content effects of the various independent variables. Recall did not differ significantly between the violent and nonviolent film nights (71 *vs* 75 percent), between focal and incidental viewing nights (72 *vs* 74 percent), or between high-exposure and low-exposure subjects (72 *vs* 73 percent).

Dream content variables. Table 2 presents a summary of all dream variable means for the two levels of each of the three independent variables (violent film *vs* nonviolent film, focal *vs* incidental involvement, low- *vs* high-exposure), along with an evaluation of the differences by F ratios or t values. No dream variable showed any significant ($p < .05$) main effect of the film manipulation. This consistent finding of "no difference" is quite striking, in view of the relatively large difference previously noted in the level of manifest violence in the two films.

Several related measures of dream hostility (Physical Aggression, Attack Social-Dreamer, and Hostility Need) showed a significant main effect of the involvement variable, with greater hostility scored following focal viewing, irrespective of the film so viewed. The significant involvement main effect for Unfavorable Outcome-Other probably also reflects this same enhancement in hostility, since the main justification for scoring Unfavorable Outcome for a character is the *receipt* of an act scored Attack Social in its *initiation.* All three of the variables which discriminate initiator from recipient thus seem to agree that the effect of focal involvement in viewing televised fantasy is a relative mobilization of hostile feelings ascribed to the self and turned toward others.

It is puzzling that this relatively greater hostility is unrelated to the film viewed. Neither main effects nor interactions were observed linking the violent or nonviolent film variable with any dimension of dream hostility. That an effect often observed in other laboratory research solely for violent stimuli could be duplicated here by a manifestly nonviolent stimulus as well as by a manifestly violent one would seem to argue for an array of well-selected control films before hostility-enhancing effects are attributed to a violent stimulus *per se.*

It is possible, of course, that the nonviolent film stimulated violent fantasies among its viewers only under focal viewing. We have no way of directly testing that possibility: were it true, however, the nonviolent film might have been expected to promote such stimulation more often in subjects with a predilection to violent fantasy than in subjects without manifest interest in such fantasy. No such interaction was observed.

Is one to conclude, then, that intense presleep viewing of any televised fantasy, violent or not, elicits hostile impulses? Such a conclusion seems premature on several grounds. It is not clear that the direction of the effect is properly stated; perhaps incidental viewing leads to a *suppression* of hostility. The hypothesis is not consistent with evidence

Table 2: Summary of main effects for dream content variables

Variable	Violent vs nonviolent			Focal vs incidental			Low vs high exposure		
	$\overline{X}$	$\overline{X}$	F[1]	$\overline{X}$	$\overline{X}$	F[1]	$\overline{X}$	$\overline{X}$	t[2]
Intensity									
Dreamlike fantasy	3.693	3.835	–	3.843	3.865	1.840	3.571	3.962	1.037
Word count	91.225	81.800	4.010	88.425	84.600	–	80.200	92.825	1.076
Hedonic quality									
Hedonic tone	3.929	3.944	–	4.006	3.866	–	3.891	3.987	.417
Unfav. out., dreamer	.256	.225	–	.242	.240	–	.209	.277	.883
Unfav. out., other	.266	.294	–	.346	.215	4.952[3]	.223	.341	1.533
Unfav. out., either	.386	.379	–	.423	.342	1.818	.329	.440	1.168
Fav. out., dreamer	.293	.321	–	.331	.283	–	.250	.369	1.676
Fav. out., other	.331	.383	–	.360	.354	–	.304	.414	1.549
Fav. out., either	.406	.433	–	.435	.404	–	.354	.490	1.766
Hostility									
Physical aggression	.877	.965	–	1.182	.660	7.018[3]	.782	1.063	.983
Attack social, dreamer	.160	.092	1.333	.202	.050	7.111[3]	.101	.154	.964
Attack social, other	.227	.300	1.112	.331	.196	3.626	.204	.327	1.952
Attack social, either	.279	.308	–	.356	.231	2.585	.250	.341	1.282
Approach social, dreamer	.275	.369	1.536	.337	.306	–	.292	.356	.831
Approach social, other	.300	.335	–	.339	.296	–	.269	.372	1.451
Approach social, either	.379	.421	–	.414	.385	–	.360	.444	1.091
Hostility need	.360	.362	–	.500	.221	7.932[3]	.272	.453	1.392
Hostility press	.818	.794	–	.997	.616	2.911	.631	.958	1.580
Defensiveness									
Total anxiety	1.399	1.305	–	1.376	1.327	1.163	1.370	1.337	.116
Death anxiety	.378	.390	–	.476	.291	1.842	.370	.399	.206
Mutilation anxiety	.338	.314	–	.283	.369	1.026	.260	.394	1.000
Guilt anxiety	.232	.244	–	.246	.230	1.087	.221	.257	.327
Rep. of setting	2.307	2.306	–	2.323	2.289	–	2.354	2.262	.561
Rep. of characters	2.295	2.272	–	2.327	2.240	–	2.465	2.106	2.546[3]
Guilt	.430	.450	–	.475	.405	–	.357	.528	1.402

[1] Blank entries signify an $F \leq 1$. At 1/34 df, $F_{.05} = 4.13$

[2] At 38 df, $t_{.05}$ (two-tail) = 2.025

[3] $p < .05$

from the other two film-dream studies, in one of which a more intensely viewed film was associated with *reduced* levels of dream hostility. It is not well supported by evidence from other media studies with waking dependent variables, and it risks confounding the particular mechanics of the viewing manipulations in the present study with more general phenomena of viewing behavior.

Focal viewing in the present study was promoted by techniques which also led to a subject's enclosure in a small, darkened room in a somewhat unfamiliar setting (and which may have been "scary" to subjects and have generated aggressive fantasies as a defense) and to a greater curtailment of his freedom to choose presleep activity, relative to the incidental condition (thus, perhaps, generating frustration and aggression). Incidental "viewing," on the other hand, was most often not viewing at all but participation in another activity. Perhaps some aspect of that nontelevision behavior can explain the hostility data. For example, it is possible that in working the Soma cube puzzle, subjects worked through feelings of hostility sufficiently well to effect a reduction in their subsequent dream hostility. Cross-validation of the present findings with different, and less confounding, techniques for eliciting desired levels of viewing behavior thus seems to be required before any generalizations can be drawn from the present data regarding the relationship of viewing intensity to dream hostility.

Only one significant main effect was noted for the exposure variable. One of the measures of "defensiveness" in dream content, the Foulkes-Shepherd character distortion scale, was scored at lower values for high-exposure subjects than for low-exposure subjects. This finding is consistent with the suggestion of some earlier research (Schramm et al., 1961; Feshbach and Singer, 1971) that children who prefer fantasy violence are lower in aggression anxiety than are their peers who prefer other media content. It thus is also consistent with a continuity hypothesis of waking-sleeping correlation. It should be noted, however, that no consistent trend was observed for low- *vs* high-exposure subjects on defensiveness measures other than Representativeness of Characterization.

Table 3 indicates the F ratios for all first-order interactions (no triple interaction was significant) for all dream variables. All significant ($p < .05$) interactions included the involvement variable as one factor. One of the measures of dream defensiveness (Representativeness of Setting) showed a significant involvement-exposure interaction: high-exposure, focal, $\overline{X} = 1.954$; high-exposure, incidental, $\overline{X} = 2.276$; low-exposure, focal, $\overline{X} = 2.583$; low-exposure, incidental, $\overline{X} = 2.158$. With focal viewing, then, by this measure, high-exposure subjects' defenses relaxed while the defensiveness of low-exposure subjects was intensified. A similar, though nonsignificant, pattern was found for the Representativeness of Characterization scale.

Table 3: Summary table of F-ratios: First-order interactions for dream content variables

	Exposure-Film	Exposure-Involvement	Film-Involvement
Variable	F[1]	F[1]	F[1]
Intensity			
Dreamlike fantasy	1.698	1.840	4.558[2]
Word count	1.864	–	6.255[2]
Hedonic quality			
Hedonic tone	–	–	5.696[2]
Unfavorable outcome, dreamer	2.159	1.363	1.257
Unfavorable outcome, other	–	–	1.012
Unfavorable outcome, either	–	–	–
Favorable outcome, dreamer	–	2.680	–
Favorable outcome, other	–	1.331	–
Favorable outcome, either	1.187	2.117	–
Hostility			
Physical aggression	1.709	–	–
Attack social, dreamer	–	2.759	–
Attack social, other	–	–	3.465
Attack social, either	–	–	3.083
Approach social, dreamer	–	–	–
Approach social, other	–	2.329	2.979
Approach social, either	–	1.802	1.445
Hostility need	–	1.759	–
Hostility press	3.345	–	2.859
Defensiveness			
Total anxiety	–	–	–
Death anxiety	3.242	–	1.579
Mutilation anxiety	2.866	–	1.281
Guilt anxiety	3.257	–	–
Rep. of setting	–	4.236[2]	–
Rep. of characters	–	4.011	–
Guilt	–	–	1.964

[1] Blank entries signify an $F \leq 1$. At 1/34 df, $F_{.05} = 4.13$
[2] $p < .05$

The three other significant first-order interactions included the film variable. As already noted, none was in the area of dream hostility. Both purported measures of general dream intensity demonstrated a significant film-involvement interaction. For Dreamlike Fantasy ratings, intensity was relatively low for violent-focal ($\overline{X} = 3.630$) and nonviolent-incidental ($\overline{X} = 3.756$) and relatively high for nonviolent-focal ($\overline{X} = 4.288$) and violent-incidental ($\overline{X} = 4.274$). For word counts, this pattern was reversed: violent-focal ($\overline{X} = 112.5$) and nonviolent-incidental ($\overline{X} = 87.9$) were relatively high, while nonviolent-focal ($\overline{X} = 78.2$) and violent-incidental ($\overline{X} = 88.9$) were relatively low. Taking subjects' two-night averages of word count and Dreamlike Fantasy ratings, a Pearson r was calculated for these two variables across the 40 subjects. The resulting correlation was positive, but not statistically significant ($r = .22$). It thus appears that the two "general dream intensity" measures are actually measuring different things. At present, the best that can be said is simply that in some respects (hallucinatory quality, bizarreness, visualization), violent-focal and nonviolent-incidental dreams are less dreamlike, while in another respect (word count) more dreamlike, than dreams collected in other conditions.

The third film-involvement interaction was for global ratings of Hedonic Tone. Unpleasantness was greater in dreams following the violent film when seen focally rather than incidentally ($\bar{X} = 4.217$ *vs* 3.624) and following the nonviolent film when seen incidentally rather than focally ($\bar{X} = 4.197$ *vs* 3.831). Although it is not difficult to account for the greater elicitation of unpleasant affect with focal viewing of the violent film, it is hard to see how that explanation also could be made to fit the equal degree of unpleasantness observed in dreams after incidental viewing of a more manifestly pleasant film. It also should be noted that film-involvement interactions for six other variables relating to the hedonic quality of dream experience were not statistically significant.

More generally, in view of the large number of comparisons made (25 dream variables, each tested for three main effects, three first-order interactions, and one triple interaction), a number of "significant" results *were to be expected by chance in the present study. Since the 25 variables* are presumed not to be totally independent of one another, the exact number so expected is not readily determinable. However, there were only five main effects and four first-order interactions observed to be "significant" at $p < .05$, and this pattern of results seems quite consistent with an overall null hypothesis for the independent variables studied.

Within such a meager harvest of significant results, moreover, one has reason to expect intrastudy consistency for conceptually similar dependent variables if some interpretive significance is to be attached to observed findings. The main effect of involvement on dream hostility shows such consistency, but it, unfortunately, is open to several different interpretations, only some of which link hostility in dreams with fantasy viewing *per se*. The interaction effect of exposure and involvement on dream defensiveness also partially meets this consistency criterion. High-exposure subjects seem to become more relaxed in their defensive maneuvers when they get highly involved in film viewing, while low-exposure subjects resort to more defensive disguise in their dream content. We interpret the effect as "relaxation" of defensiveness rather than a "breakdown" of defensive structures, since there is no attendant increase in unpleasant affect or hostile impulse.

Alternatively, one is likely to ascribe interpretive significance to a result, even when it is observed among many nonsignificant findings, when that result is in agreement with prior research, with some general psychological theory, or with common sense expectation. Only the finding of less character distortion for the dreams of high-exposure subjects even approximately meets this criterion.

Overall, then, some "statistically significant" results were observed in dream content variables, but it is doubtful that much interpretive significance should be attached to them at this time. The finding of hostility mobilization by the focal viewing condition deserves restudy, as does the lesser distortion in the nocturnal fantasy of high-exposure subjects, particularly under focal involvement. Otherwise, the results we have obtained are best viewed as indicating no probable effects for the independent variables—including, most clearly, the presleep viewing of a violent *vs* a nonviolent film.

DISCUSSION

Film variable

The findings of the present study indicate that the violent film did not have any systematic effect on dream hostility, anxiety, guilt, hedonic tone, or overall vividness and intensity. They thus are in direct agreement with the results of neither of the earlier two studies relating media violence to dream content (Foulkes and Rechtschaffen, 1964; Foulkes et al., 1967). The former study, with a young adult sample, found an increase in general intensity; the latter, with preadolescents, found a decrease in intensity, aggression, "bad dreams," and guilt following exposure to violent films.

In addition to attempting to replicate earlier findings on the effects of violent *vs* nonviolent films, the present study also sought to extend the understanding of such effects by determining their dependence on the quality of viewing behavior shown to film stimuli and a subject's prior history of exposure to violent media stimuli. No significant interactions were observed between the prior-exposure variable and the violent *vs* nonviolent film variable. Three statistically significant interactions were observed between the film variable and the experimental manipulation of viewing involvement. These three findings were not part of any consistent pattern, however, and were not corroborated by other dream ratings presumed to measure similar variables as those for which significance was observed. Thus, we have interpreted the three statistically significant interactions which were observed as not meriting serious consideration at this time for generalization beyond the confines of this experiment and as probably having been due to chance, given the large number of comparisons from which only they emerged as film-related effects.

The discrepancy between the present lack of positive results for the film manipulation and the findings of earlier studies deserves some comment. Our remarks are grouped into two areas: comparisons of methodology and comparisons of results.

Methodology. We stressed earlier that details of methodology are of critical importance in evaluating generalizations from laboratory experiments to those social contexts which the laboratory manipulations were intended to simulate. This has been our justification for critically examining earlier studies and for a relatively lengthy presentation of our own methods. In considering the implications of the present study relative to those of the two earlier film-dream studies, then, it will be helpful to make a comparative analysis of their methodology. Table 4 lists a number of characteristics of a film-dream study which would seem to enhance generalization from laboratory to everyday social reality.

Table 4: Three film-dream studies: comparative methodology and results

Method	Foulkes and Rechtschaffen (1964)	Foulkes et al. (1967)	Present study
Subjects:			
typical media consumers?	0	X	X
objects of social concern?	0	X	X
broad community sample?	0	0	X
Films:			
representative television product?	X	0	X
seen complete with commercials?	0	0	X
control matched for series?	X	0	X
equivalent interest value?	0	0	X
representative levels of violence?	X	X	?
Viewing situation:			
naturalistic presentation?	0	0	X
rationale for viewing?	0	X	X
credibility of rationale?	—	?	X
long-term viewing?	0	0	0
Dependent variables:			
naïve interviewer?	0	0	X
postsleep observations?	0	0	0
long-term evaluation?	0	0	0
Effects of Violent Film			
No increase in sleep disturbance	X	X	X
No increase in hostility	X	X	X
No increase in anxiety or guilt	X	X	X
No increase in negative feeling tone	X	X	X

With respect to *subjects*, only the two child studies employed subjects who typically are heavy consumers of televised fantasy and about whose viewing most social concern seems to be expressed. When we compare the sampling of children in these two studies, it appears that the present study employed the more representative technique: advertisement in the community's newspaper as compared to the earlier study's advertisement in the campus newspaper. Social class data were not obtained systematically in the earlier study, but it is known that at least two-thirds of the sample consisted of sons of university faculty or students, compared to only 30 percent in the current study. While even the latter figure obviously is quite high for representative national sampling, it is not so out of line for representative community sampling in a small university town. The boys in the present study came from all social classes and all sections of town. The mere size of the volunteering group (N = 105) for a narrow age band in a community of approximately 20,000 suggests a relatively broad sampling.

Films were representative media products only in the young adult study and the present one. The film employed in the 1967 study, although available commercially in the form in which it was shown, was in fact a rather inept pastiche of scenes from a feature-length film, with much action but little characterization or thematic coherence. Only in the present study was a television program seen as it ordinarily is, complete with commercials. Only in the young adult study and the present study was the control film from the same television series as the violent film, thus offering some comparability of characterization, setting, and plot elements between the two stimuli. The baseball control film of the 1967 study might have had properties of its own leading to heightened dream aggression, rather than the experimental film producing decreased aggressive dream content. As already has been noted, experimental and control films in the two earlier studies differed in interest values for the subjects. In the present study, there was no indication that one film was more interesting to subjects than the other. During the first series of experimental nights, there was a nonsignificant trend for behavior ratings of attentiveness under incidental involvement to be higher for the violent film; in the second series, the tendency was for eye contact with the stimulus on incidental involvement nights to be higher for the nonviolent film. Overall, then, the two film stimuli of the present study seem to have been roughly comparable in their general interest to subjects.

One question about the violent stimulus in the present study is whether it was sufficiently violent to be representative of television programs which have been the object of social concern. The use of a "children's western" may have been somewhat unfortunate, in that most such concern seems to attach to the situation in which children watch program primarily designed for adults. There may be a sufficiently clear formula for children's westerns which renders the audience less susceptible to effects there than to the more unpredictable (to a child, at least) world of adult fantasy violence. On an absolute basis, the violent film in the present study was much less explicit than the film used in the young adult study. In that film, for example, a doctor ministered to the needs of an escaped convict and then, after having begged for his life, was shot in the stomach with his agony plainly portrayed on the screen. As a result of increasing public concern over the effects of televised violence, such a scene would not likely be shown on television today.

Yet the violent stimulus of the present study *was* a manifestly violent episode of a series currently felt by the general public (Greenberg and

Gordon, 1971) to be among television's 20 most violent series. In portraying a boy as the victim of kidnapping, verbal abuse, and physical attack, moreover, it might be considered potentially more threatening to young subjects than the content of most "adult" violent fantasy. In any event, if not representative of levels of violence that have been shown on television in the past, the violence of the experimental stimulus in the present study probably was sufficiently representative of televised violence *circa* 1969-70 to permit some generalization to that season's violent programs. At the very least, the results of the present study cast doubt on the generalization some researchers (e.g. Goranson, 1970) have drawn from laboratory studies to the effect that "sanitizing" violent media fare may make it more harmful.

With regard to the *viewing situation*, our criticism of much of the research of the 1960s has been that reasons and mechanisms for delivering media stimuli have been such as to put laboratory viewing behavior at some psychological distance from everyday viewing behavior. How do the three film-dream studies stand on this score? The young adult study of Foulkes and Rechtschaffen cannot be faulted for its deceptions: subjects there knew that the study was an attempt to trace the dream effects of the stimuli they were viewing. Yet this degree of frankness unfortunately does not make for a natural climate for presleep film viewing, nor was the situation of watching the television films projected on a movie screen in the presence of the projectionist a contribution to representative viewing behavior. The 1967 child study employed the same mechanism for showing films. In that study, however, subjects were given a deceptive cover story about why they were being shown movies. It was not a very clever story, and the experimenters received indications that at least some of the subjects saw through it.

In the present study, a television program was transmitted by a real and identifiable television station to an ordinary television receiver. Thus the mechanism for delivering television films to subjects was entirely natural. Rather than telling subjects they were investigating the effects of television programs on dreams, the experimenters gave a somewhat deceptive rationale to explain the presence of the television receivers in the bedrooms and the nature and conditions of the viewing behavior subjects were permitted. As already has been noted, this rationale was in accord with several facts known to, or experienced by, the subjects (the change in laboratories, the other children's sleep studies engaged in by the same laboratory), was not on its face implausible, and did not seem to arouse the suspicion of the relatively inexperienced and unsophisticated subject sample. Because the boys were "subjects" in an "experiment," however, and not simply themselves in their everyday world, the representativeness of their viewing behavior in the laboratory requires further discussion.

In spite of the systematically different degree to which we facilitated or impeded viewing, it might fairly be said that the subjects' motivation for watching the two television programs was spontaneous. That is, they watched television more because they wanted to than because we may have asked them to. As we have already noted, a frequent question subjects asked upon arriving at the laboratory and inspecting their rooms was if they could watch television. The experimenters thus became responsible more for the timing of the viewing and the content viewed than for the fact of viewing itself. The most deprivation of freedom of choice subjects faced was in the focal condition, where they were asked to remain awake and watch television. Everyday experience with preadolescent boys indicates that these activities generally are relatively preferred to going to sleep at 9 p.m. and watching no television, and that these activities have a generally high absolute attraction. In the incidental condition, the effect of the manipulation was essentially to give subjects genuine freedom of choice with respect to presleep activities, except that, again, sleep was prohibited.

The justification presented to subjects for their having to watch the particular program they did was that we were not yet connected by antenna or cable to any other station. This statement, as subjects could and sometimes did verify by checking the different channels, was entirely accurate. As several of their own comments suggested, the subjects found themselves in a situation not unlike that of sleeping away from home in a motel. Consistent with the analogy was their seeming acceptance of the conditions of viewing the laboratory afforded: as when they might be in a motel, it was possible for them to watch only those programs that could be received on the sets provided. As this analogy also implies, it is our feeling that the subjects' viewing behavior was general unselfconscious and that, if it was performed in a slightly different than homelike context, it still was not unlike some natural viewing situations and was approached by subjects with the same attitudes as most everyday contacts with television.

The study clearly was not naturalistic; the boys knew that they were subjects—in a sleep experiment. Perhaps this is one direction in which media experiments might move to bridge the gap between laboratory situations and everyday media behavior. If the media exposure can be made to appear as a background or incidental feature of a study with another clearly articulated rationale, it may become less tied up with the unique features of laboratory interaction which seem to compromise generalizations to the "real world." Conversely, of course, it may also become contaminated by adverse properties of the "other experiment" in which it is enmeshed. For this reason, it is appropriate to say something about the way in which a laboratory sleep experiment is perceived by preadolescent boys.

In our experience, electrophysiological sleep and dream studies have elicited considerable interest and cooperation from preadolescents serving for several nights in the laboratory. We are sufficiently experienced in studies with subjects of this age range to feel comfortable with them. The social atmosphere of the laboratory is friendly and informal—not on a "medical model." The new laboratory was designed with children's sleep studies specifically in mind. It contains a chalkboard and colored chalk for artistic endeavors and several attractive toys for play. (It also generally contains a number of children's books and magazines, but they were removed prior to the start of the present study to avoid uncontrolled media mixtures.) Subjects were able to engage in some play behavior with their peer companions prior to retiring, with much attention given to a "Krazy Kar" in which they could wheel around the large laboratory play area at high speeds. The laboratory is relatively uncluttered with technical apparatus. What machinery is present is in one partly enclosed area and seems more to fascinate than to frighten subjects. (It is of some interest that laboratory apprehension seems far more common among college students we have run than among grade-school children.)

In the present study, the aura of a new laboratory, the limited gadgetry, the easy social atmosphere, the television receivers, the toys, and the relatively generous payment all seemed to contribute to making the sleep experiment something other than a stressful test. The laboratory seemed to most boys a pretty interesting place to be, a place where they could behave in a relatively relaxed and unselfconscious way. We cannot deny that some subjects gave minimal signs of discomfort and apprehension; we merely wish to point out that this was not the predominant impression and that one should not infer from the physical dissimilarity to home or from one's own adult judgments of the situation that the sleep laboratory elicited an entirely unnatural set of responses from subjects.

One very real limitation shared by all three studies in the series under evaluation is that they are limited to testing the effects of a single exposure to a single program. The hazards in generalization to long-term effects of repeated exposure already have been noted (e.g., the quotation above by Feshbach, 1963). The discrepancy between results of the Feshbach-Singer (1971) long-term study and the traditional short-term laboratory studies of the 1960s suggests the shortcomings of the latter in any meaningful extrapolation to long-range effects. The present study, however, is not totally lacking in data which may be relevant to the question of long-range effects. Comparisons of questionnaire responses and dreams of high- and low-exposure subjects, although limited in that they cannot suggest a direction of causality, are germane to the question.

Dependent variable considerations also suggest common limitations of all three studies. Some sort of postsleep test of aggression would have strengthened one's confidence in the meaning of their dream data and provided evidence about whether dream expression substitutes for, stimulates, or is unrelated to subsequent waking expression. Long-run evaluations of the effects of film exposure also were lacking at any level —dream or waking. In the absence of short-term effects, this may not be a critical flaw in the three experiments under consideration, although some sort of "sleeper" effect conceivably might exist. Still, it would be nothing short of amazing to find significant long-range effects of viewing a single, typical television program: that kind of evaluation seems more appropriate to the study of an extended period of exposure to violent fantasy. Real advances in research on socially relevant effects of fantasy violence probably cannot be made by one-shot experiments or crash programs; they await a commitment—both by researchers to execute and by society to underwrite—long-term studies of exposure and effects.

One technical flaw in the design of the prior two studies not repeated here was the use of an interviewer who knew the presleep film condition. We do not believe that this difference in method explains the discrepancies in results, because the two prior studies shared the same interviewer but, contrary to anticipation at the time, not the same results. Still, the greater confidence must attach to the results of the present study in this area.

Overall, then, none of these three laboratory experiments permits any easy generalization to the social context to which they addressed themselves. No such experiment or series of experiments ever does. On a relative basis, however, we have tried to justify the position that the methodology of the present study was the most appropriate to questions generating media studies and the most naturalistic.

Results. The lower section of Table 4 is intended to convey that the seeming inconsistency of film effects among the three studies may be more apparent than real in several quite important respects. None of the three studies provides any support for the hypothesis that presleep viewing of a typical media portrayal of violence is associated with disturbance of subsequent sleep. Even more significantly, none of the three studies provides any support for the hypotheses that viewing a violent film *per se* stimulates hostility, anxiety, guilt, or an unpleasant feeling tone.

The objection may be raised that dream content studies cannot be used to refute the results of other laboratory studies in which there are waking dependent variables of more manifest relevance to socially significant behavior. We do not intend to suggest that dream variables are "better" than other variables for assessing the effects of televised violence, nor that they can substitute for them. We reiterate, however, that the waking dependent variables of many other laboratory studies (including the two we have reviewed above in detail) also are not obviously related to socially significant aggressive behavior, and that there are several respects in which dream variables are less susceptible to contamination limiting their generalization than are the waking tests which many other studies have employed. That dream data are more irrelevant than the waking measures of aggression typically employed in laboratory experimentation cannot be taken as a foregone conclusion. This issue is still an open one. It does seem, however, that when two new lines of evidence—namely the Feshbach and Singer (1971) field study and these dream investigations—do not support the results of the traditional laboratory studies, then the generality of the laboratory studies has become somewhat questionable.

Conceptually, it would not seem surprising that media violence should have few significant harmful effects. People seem to enjoy media violence, as they have from time immemorial presumably enjoyed hearing violent folk tales and myths of nonmass media. The pervasiveness of violent fantasy suggests that it fulfills at least some integrative functions for individuals and for societies. While there may be latent harmful consequences of exposure to such material, the persistence of the human liking for violent fantasy suggests that they must be at least partially outweighed by other, more useful, consequences. Among the latter, as Feshbach and Singer have suggested, may be the development of a "cognitive control" system for coping with one's own impulses.

Why do children like media violence? Himmelweit et al. (1958) found that children liked to be frightened a little by what they saw. Children discriminated between "exciting" programs, which they liked, and "frightening" ones, which they did not. This discrimination was made in relation to the child's ability to cope with emotions aroused by program content. What was frightening to children in one age group, for example, might be merely exciting to older children. To the youngest children, westerns were frightening; to early adolescents, they generally were not, while the less stylized violence of crime shows still was. As children mature, their taste changes, presumably in relation to changing developmental needs and competencies. There may be a set of hurdles in television violence—animal cartoons, simple westerns, adult westerns and crime programs, supernatural terror shows—which, if taken out of sequence, can be harmful, but which, if taken as children generally choose to take them, serve adaptive functions in relation to the current stage of ego development. The word "hurdle" is used advisedly, for children seem impelled to master each new level, although initially it may frighten them. Mastery of such material might well correspond to personal mastery and growth, and not be causally unrelated to these favorable developmental trends. In any event, it is easy to lose sight of the fact that children do like violent media fantasy, and that behind this liking may lie something less pathognomic than a blind drive to self-destruction or egoism.

The developmental hurdle argument also implies that violent fantasy may, at points of transition to new and "frightening" material, stimulate "bad" impulses or feelings such as aggression or guilt. If they are way stations on the path to better control and greater maturity, however, even these effects must be evaluated in context. The control of impulsivity is not likely to be achieved by its total suppression, nor are all unpleasant affects or aggressive behaviors to be immediately condemned as personally and socially undesirable. Partial elicitation of such phenomena, by the mass media or other means, may be necessary to the full elaboration of personal and social systems of control over impulsive behavior. Limits, of course, are part of the development of any dynamic regulatory system, and one wonders if the mass media are not being subjected to excessive blame for a breakdown in those family and other structures that heretofore may have kept the problem of impulse regulation in better dynamic balance.

The force of this argument is that even if results indicated increased aggression, anxiety, etc., following the exposure of children to violent films, such effects could be evaluated as intrinsically undesirable only if one holds to the theory that these impulses and feelings are themselves inherently undesirable. If one rejects that theory, however, any stimulating effects must be evaluated in the context of the total development of the individual's system for regulating feelings and fantasies and in the context of the total system of social control over impulsivity. Any faulty outcomes in the development of either system may be attributable not so much to the stimulation of feelings or fantasies by whatever means as to the failure to make adequate provision for their realistic regulation.

Involvement variable

In retrospect, there are several ways in which the involvement manipulation may have been too "successful." It created such large mean differences in viewing behavior between focal and incidental involvement that the differentiation did not adequately represent either levels of viewing involvement generally shown by children to televised fantasy or viewing behavior shown by subjects in the two earlier film-dream studies.

In terms of generalization to everyday viewing behavior, the nearest parallel to our incidental condition is the child sitting on the living room floor engaged in some activity while the television receiver is turned on. He looks at it from time to time, but his attention is mainly elsewhere. This is not a viewing situation which has elicited much social concern. Because of the magnitude of difference between viewing on focal and incidental nights, we are unable to say anything about the more socially relevant comparison of children held in rapt attention by a media stimulus versus children engaged in "ordinary" viewing behavior.

The involvement manipulation was intended to test a *post hoc* hypothesis suggested by the two earlier film-dream experiments. In both of them, the more intensely attended-to stimulus was associated with less intense subsequent dreams. But the degree of difference in viewing behavior between our focal and incidental conditions, and the means used to generate this difference, render our results unsuitable for comparison with those of the earlier studies. In those studies, subjects had no meaningful alternative to watching the film: they watched with more or with less enthusiasm, but they watched. In the present study, the incidental condition allowed for much non-viewing and so is not comparable to the "less involved" viewing of earlier studies. The reason for the difference, of course, was the presence in our incidental condition of meaningful alternatives to viewing.

More appropriate than the overall incidental *vs* focal involvement results to the hypothesis suggested by the earlier studies may be the data on involvement actually elicited on focal nights. Here it is known that subjects did view the films in approximately the same manner as in the earlier studies; there were no experimentally provided distractions from high levels of viewing behavior. Within the focal viewing condition, there was no evidence of systematic difference in attentiveness to the violent *vs* the nonviolent film. However, there were differences in dream intensity following the two films when they were seen under focal involvement (higher dreamlike fantasy but lower word counts following the nonviolent film). In this sense, the data, although not self-consistent, do not seem to be confirmatory of the *post hoc* hypothesis that differences in viewing interest lead to differences in dream vividness.

Given the low degree of viewing behavior elicited under the incidental condition of the present study, might not the incidental *vs* focal main effects be used to test another hypothesis—namely, that there are different effects of viewing filmed fantasy *vs* not viewing filmed fantasy? Intersubject variability in viewing behavior actually elicited under the incidental manipulation suggests that incidental "viewing" may have been essentially nonviewing for some subjects, but that it was margin-of-consciousness viewing for others and sporadic but sometimes relatively involved viewing for still others. Thus, if the relatively low mean for actual viewing behavior in the incidental condition represents too high a degree of "success" in the manipulation of viewing, the variability around that mean represents another and more manifest form of failure —the failure to create a psychologically homogeneous state within a given level of the involvement treatment. Such inhomogeneity renders interpretations of the effect of incidental viewing rather equivocal. On the one hand, one might be dealing almost with a no-film control condition; on the other hand, there might be a sufficient degree of stimulation in incidental viewing to warrant its consideration within the range of genuine viewing behavior.

To these caveats concerning the interpretation of involvement main effects, of course, must also be added the previously noted confounds between the particular situations used to create focal *vs* incidental involvement and the level of viewing actually observed in these situations. The only consistent main effect of the present study was greater dream hostility following focal than following incidental involvement, but the presence of such confounds, plus the lack of any other clear rationale for the observed pattern of results, has seriously beclouded the meaning of this one instance of intrastudy consistency in the main effect data.

Acknowledging these confounds brings one back to the observation that viewing behavior was rather too "successfully" manipulated in the present study. The confounds all derive from an attempt to create an incidental viewing condition that would contrast sharply with the more typical focal viewing situation. Thus, it was thought necessary to give incidental condition subjects genuine freedom of choice over their presleep activities. In contrast to focal condition subjects, they had a range of attractive presleep activities from which to choose and no restriction on their choice. However, measurement operations were not extended accordingly. "Involvement" was assessed only in relation to film stimuli. Thus there was neither homogeneity in what incidental subjects did nor systematic assessment of all of the various possibilities offered to them. Consequently, there is no more evidence to allow an evaluation of the general hypothesis that presleep variations in enthusiasm or interest in any sort of activity will be associated with variations in the intensity of subsequent dreams than there is to allow clear-cut evaluation of other hypotheses relating more specifically to film involvement.

Prior exposure variable

This variable was defined in an "applied" manner. Rather than choosing some behavioral science construct like social class or intelligence for discriminating the subgroups (and then having to qualify one's generalizations with observations such as "Lower-class children *tend* to view more televised violence, and they also tend to. . ."), there need be no qualification concerning the differences between the two groups in this study in terms of their viewing behavior. There was no overlap. From the point of view of scientific understanding, however, it seems important to try to relate the real-world variable of high *vs* low violence exposure to the traditional variables of psychological and sociological inquiry. Our attempts to do this met with only limited success, but the available questionnaire data also were quite limited. It does seem, however, that when one begins with exposure and works back to these traditional scientific variables, the gap between the domains of social concern and scientific theory looms very large indeed. For example, "addicts" to television violence can come from any social class or any level of intellectual achievement. It remains to be seen whether the strategy of "pure" research, to begin with its own established variables, can meaningfully bridge this gap at present, or whether it might not be preferable to start with the applied variable of exposure, if socially relevant generalizations are needed.

Only one main effect and one interaction were significant for the exposure variable. Less character distortion was found for subjects with a high interest in, and high exposure to, violent media fantasy. Setting distortion showed an exposure-involvement interaction (as also did character distortion, but not significantly), with lowest levels of distortion found for focal viewing by high-exposure subjects and for incidental viewing by low-exposure subjects. In each of these cases, dream defensiveness was reduced when subjects viewed media stimuli in a manner which might have been most familiar to them: intensely for subjects with a history of high exposure to television fantasy, and sporadically for subjects with a history of low exposure to television fantasy. To state it conversely: dream censorship increased when high-exposure subjects were unable, in a particular situation, to maintain high exposure and when low-exposure subjects were unable to maintain low exposure.

Although the main effect is comprehensible in terms of a continuity of defensive style, and although the interaction makes sense in terms of the conditions underlying dream censorship, the failure to find other significant differences or interactions pertaining to the exposure variable seems to indicate an absence of possible major effects for long-term high exposure to televised violence, at least for boys in the age group studied. It is possible, of course, that such effects might become evident later in adolescence, when viewing behavior is not so normative for an entire age group. In that case, however, the question would remain whether the continued viewing of some adolescents would be responsible for personal or social dysfunction or whether personal-social problems might not be responsible for their continued high rate of viewing fantasy violence. The alternatives, of course, are not mutually contradictory.

The findings of Feshbach and Singer (1971) seem to indicate that, while boys with a preference for aggressive activities tend to take a higher interest in aggressive media fantasy, boys generally are rendered less, not more, aggressive by viewing such fantasy. In any event, it can be stated that if habitual high exposure to televised violence does have any deleterious psychological consequences, they did not show up in the present study of male preadolescents, a finding made more striking by virtue of the extremity of the two groups, viewing habits.

The lack of dream differences between high- and low-exposure subjects may be compared to the results of several earlier studies in which pathology, rather than viewing of violent fantasy, was the independent variable (Foulkes and Rechtschaffen, 1964; Pivik and Foulkes, 1966; Foulkes et al., 1969). In these studies, high-pathology subjects have reported dreams of increased general intensity and unpleasant affect. In these respects, the high-exposure group of the present study does not seem to be comparable to high-pathology groups previously studied.

This in turn suggests that personality dysfunction is not a dimension which significantly discriminates high-exposure subjects from their low-exposure peers.

SUMMARY

The present study attempted to clarify and extend the results of two earlier studies relating laboratory-collected rapid eye movement (REM) dreams to presleep viewing of violent *vs* nonviolent films. Following an adaptation night in a sleep laboratory, 40 male preadolescents were studied for two additional experimental nights. On one night, they saw a nonviolent episode from a children's western television series; on the other, they saw a violent episode from the same series. Films shown on ordinary television receivers were an apparently incidental feature of the laboratory setting. For some subjects, both films were viewed under experimentally manipulated focal involvement; for others, both were viewed under experimentally manipulated incidental involvement; for still others, one film was viewed under each of the two involvement conditions. Direct behavioral observation and closed-circuit television monitoring established that different viewing behaviors were in fact elicited in the two involvement conditions. Twenty of the subjects had indicated a history of much prior exposure to televised violence, while 20 had reported minimal prior exposure to it. The two subject groups were the extremes on the dimension of habitual exposure to televised violence among 105 volunteers for the experiment. Order effects were balanced across the film and involvement variables within both groups. The design thus was a 2^3 factorial one (nonviolent *vs* violent film, focal *vs* incidental viewing, high- *vs* low-exposure subjects), with incomplete blocking.

During the sleep period subsequent to the film manipulation, subjects were monitored by EEG and EOG recording to permit the detection of REM periods. The first four REM periods of ten minutes' duration were interrupted at that point for the retrieval of dream content by an interviewer naïve as to subjects' status on the independent variables. Recordings of these interviews subsequently were transcribed; the transcriptions were shorn of information identifying subjects or nights and then presented to judges. The judges scored the dreams for at least two separate, standard measures in each of the four areas of dream content: overall vividness, hedonic tone, hostility, and guilt-anxiety. Interrater reliability of their judgments was satisfactory.

Dreams were recalled on 73 percent of 312 experimental awakenings. While, predictably, some significant results were observed among the 175 separate comparisons subjected to statistical analysis (three main effects and four interaction terms for each of 25 dream scoring variables), they were not numerous and only occasionally found confirmation in results with conceptually related variables. Measures of dream hostility were related only to the focal-incidental manipulation of viewing behavior, and the possibility was raised that even these results might reflect the particular mechanics employed in creating differential viewing behavior in the present study rather than variations in viewing behavior *per se.* Measures of dream distortion or censorship showed some association with the exposure variable: high-exposure subjects, particularly under focal involvement, had somewhat less distortion in areas of dream characterization and setting.

No significant main effects were observed for the film variable. Two of three significant interactions involving the film condition were seemingly mutually inconsistent, and the third was unsupported by other relevant data. None of these interactions was for a hostility variable. The predominantly negative results for the film variable did not provide strong confirmation for the results of either of the two earlier studies, which found increased vividness or decreased vividness and hostility following exposure to a violent film. Methodological differences among the three studies were examined. We concluded that the procedures of the present study included the most naturalistic context for presleep film viewing and were most appropriate for generalizations to children's everyday television behavior. We also pointed out that, despite some differences in results among the three studies, none of them has demonstrated that violent films increase hostility, anxiety, guilt, or unpleasant affect in dream content or induce disturbed sleep.

FOOTNOTES

1. The research on which this report is based was conducted pursuant to the terms of Contract No. HSM 42-70-62 with the U. S. Department of Health, Education, and Welfare.
2. We are grateful to Donald A. Anderson for suggesting the format of the statistical analysis and to Donald W. Stilson for statistical consultation.
3. Although two other network affiliates are available to cable television families (one of which also is available to antenna-reception families), their programs are a multinetwork selection of the Denver offerings.
4. All comparisons between questionnaire responses of the two subject groups were evaluated by the Wilcoxon Rank-Sum Test (Wilcoxon and Wilcox, 1964).
5. *Brady Bunch* also received four votes from high-exposure subjects, and it tied for third most popular show in that group. In relation to arguments often advanced to the effect that violence is what the public and its children want, it is significant to note that this nonviolent program was able to achieve equivalent high appeal among

both violence "addicts" and those preferring nonviolent shows. More generally, the fact that a majority of favorite programs for subjects "addicted" to televised violence were not in fact violent recalls Himmelweit et al.'s (1958) finding that as many as two-thirds of the children's choices were of other than violent programs as favorites and their suggestion that violent content on television may exist out of proportion to the extent that children want (or need) it.

6. These questions were taken from unpublished materials supplied by Urie Bronfenbrenner.
7. In the case of one subject, 14 days intervened between the second and third night.
8. There were several exceptions to this generalization, due to the fact that two boys, on separate nights, missed their initial appointments for Night 3. On those two nights, only the remaining two subjects served. The two delinquent subjects spent their third night together although they had not served together previously. Overall, 42 subjects were run through the experimental series. Data from one of the aforementioned delinquents was discarded because of an unusually long temporal gap between his second and third night, while data from another subject had to be discarded when a visitor from the sponsoring organization disclosed the rationale of the study in a chance meeting with the subject's mother.
9. Robert Kelso and Terry Brubaker served as interviewers.
10. We wish to thank the National Broadcasting Company, particularly Thomas Coffin and Marjorie Shields, and Twentieth Century-Fox for their assistance in allowing us to use these films on an extended loan.
11. We are most appreciative of the cooperation received from Broadcast Services of the University of Wyoming, particularly from John McMullen, Dave Worley, and Bob Rule.
12. On only one night was there any delay (about 20 minutes) between the subjects' being put to bed and the start of the television program.
13. Parker Brothers, Inc., Salem, Massachusetts.
14. Jean Shepherd, Marti Krauss, and Robert Kelso. (Kelso did not, of course, observe subjects whom he was responsible for interviewing; his share of the interviewing assignment was completed before he participated in the behavioral observation task.)
15. Terry Brubaker, Gerry Wilcove, Edward Belvedere, Sharon Frost, and Jean Shepherd served as judges.

REFERENCES

Adler, A. *What life should mean to you.* New York: Capricorn, 1958 (orig. 1931).

Allyn, J., and Festinger, L. The effectiveness of unanticipated persuasive communications. *Journal of Abnormal and Social Psychology*, 1961, **62**, 35-40.

Argyris, C. Some unintended consequences of rigorous research. *Psychological Bulletin*, 1968, **70**, 185-97.

Aserinsky, E., and Kleitman, N. Two types of ocular motility occurring in sleep. *Journal of Applied Physiology*, 1955, **8**, 1-10.

Bailyn, L. Mass media and children: a study of exposure habits and cognitive effects. *Psychological Monographs*, 1959, **73**, no. 1 (whole no. 471).

Bandura, A. Vicarious processes: a case of no-trial learning. In Berkowitz, L. (Ed.) *Advances in experimental social psychology*, Vol. 2. New York: Academic Press, 1965.

Bandura, A., Ross, D., and Ross, S. A. Imitation of film-mediated aggressive models. *Journal of Abnormal and Social Psychology*, 1963, **66**, 3-11.

Barker, W. J. The stereotyped western story. *Psychoanalytic Quarterly*, 1955, **24**, 270-80.

Ben-Horin, P. The manifestation of some basic personality dimensions in wakefulness, fantasy, and dreams. Unpublished doctoral dissertation, University of Chicago, 1967.

Berkowitz, L. *Aggression: a social psychological analysis.* New York: McGraw-Hill, 1962.

Berkowitz, L. The effects of observing violence. *Scientific American*, 1964, 210 (no. 2), 35-41.

Berkowitz, L. Impulse, aggression and the gun. *Psychology Today*, 1968, 2 (no. 4), 19-22.

Berkowitz, L., and Rawlings, E. Effects of film violence on inhibitions against subsequent aggression. *Journal of Abnormal and Social Psychology*, 1963, **66**, 405-12.

Campbell, D. T., and Stanley, J. C. Experimental and quasi-experimental designs for research on teaching. In Gage, N.L. (Ed.) *Handbook of research on teaching*. Chicago: Rand McNally, 1963.

Dement, W., and Kleitman, N. The relation of eye movements during sleep to dream activity: an objective method for the study of dreaming. *Journal of Experimental Psychology*, 1957, **53**, 339-46.

Dysinger, W. S., and Ruckmick, C. A. *The emotional responses of children to the motion picture situation.* New York: Macmillan, 1933.

Emery, F. E. Psychological effects of the western film: a study in television viewing (I). *Human Relations*, 1959, **12**, 195-213.

Eron, L. D. Relationship of TV viewing habits and aggressive behavior in children. *Journal of Abnormal and Social Psychology*, 1963, **67**, 193-96.

Feshbach, S. The stimulating versus cathartic effects of a vicarious aggressive activity. *Journal of Abnormal and Social Psychology*, 1961, **63**, 381-85.

Feshbach, S. The effects of aggressive content in television programs upon the aggressive behavior of the audience. In Arons, L., and May, M. A. (Eds.) *Television and human behavior: tomorrow's research in mass communication.* New York: Appleton-Century-Crofts, 1963.

Feshbach, S., and Singer, R. D. *Television and aggression.* San Francisco: Jossey-Bass, 1971.

Foulkes, D. Dream reports from different stages of sleep. *Journal of Abnormal and Social Psychology*, 1962, **65**, 14-25.

Foulkes, D. *The psychology of sleep.* New York: Scribner's, 1966.

Foulkes, D. A longitudinal study of childhood dreaming. *Psychophysiology*, 1969, **6**, 266-67 (abstract).

Foulkes, D. The dreamlike fantasy scale: a rating manual. *Psychophysiology*, 1970, **7**, 335-36 (abstract).

Foulkes, D., Larson, J. D., Swanson, E. M., and Rardin, M. Two studies of childhood dreaming. *American Journal of Orthopsychiatry*, 1969, **39**, 627-43.

Foulkes, D., Pivik, T., Steadman, H. S., Spear, P. S., and Symonds, J. D. Dreams of the male child: an EEG study. *Journal of Abnormal Psychology*, 1967, **71**, 457-67.

Foulkes, D., and Rechtschaffen, A. Presleep determinants of dream content: effects of two films. *Perceptual and Motor Skills*, 1964, **19**, 983-1005.

Foulkes, D., and Shepherd, J. Manual for a scoring system for children's dreams. Laramie: Department of Psychology, University of Wyoming, 1971 (mimeo).

Freud, S. *The interpretation of dreams.* New York: Basic Books, 1956 (orig. 1900).

Goranson, R. E. Media violence and aggressive behavior: a review of experimental research. In Berkowitz, L. (Ed.) *Advances in experimental social psychology*, Vol. 5. New York: Academic Press, 1970.

Gottschalk, L. A., and Gleser, G. C. *The measurement of psychological states through the content analysis of verbal behavior.* Berkeley and Los Angeles: University of California Press, 1969.

Gottschalk, L. A., Winget, C. N., and Gleser, G. C. *Manual of instructions for using the Gottschalk-Gleser content analysis scales: anxiety, hostility, and social alienation-personal disorganization.* Berkeley and Los Angeles: University of California Press, 1969.

Greenberg, B. S., and Gordon, T. F. Perceptions of violence in TV programs: Critics and the public. *Television and social behavior:* a report to the Surgeon General, Vol. 1. Washington, D. C.: U. S. Government Printing Office, 1971.

Hauri, P. Effects of evening activity on subsequent sleep and dreams. Unpublished doctoral dissertation, University of Chicago, 1966.

Hauri, P., Sawyer, J., and Rechtschaffen, A. Dimensions of dreaming: a factored scale for rating dream reports. *Journal of Abnormal Psychology*, 1967, **72**, 16-22.

Himmelweit, H. T., Oppenheim, A. N., and Vince, P. *Television and the child.* London: Oxford University Press, 1958.

Homans, P. Puritanism revisited: an analysis of the contemporary screen-image western. *Studies in Public Communication*, 1961, **3**, 73-84.

Jung, C. J. General aspects of dream psychology. In *The structure and dynamics of the psyche.* New York: Pantheon, 1960 (orig. 1916, rev. 1948).

Kelman, H. C. Human use of human subjects: the problem of deception in social psychological experiments. *Psychological Bulletin*, 1967, **67**, 1-11.

Lazarus, R. S., Speisman, J. C., Mordkoff, A. M., and Davison, L. A. A laboratory study of psychological stress produced by a motion picture film. *Psychological Monographs*, 1962, **76**, no. 34 (whole no. 553).

Maccoby, E. E. Why do children watch television? *Public Opinion Quarterly*, 1954, **18**, 239-44.

Monroe, L. J., Rechtschaffen, A., Foulkes, D., and Jensen, J. Discriminability of REM and NREM reports. *Journal of Personality and Social Psychology*, 1965, **2**, 456-60.

Orne, M. T. On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *American Psychologist*, 1962, **17**, 776-83.

Pine, F. Incidental stimulation: a study of preconscious transformations. *Journal of Abnormal and Social Psychology*, 1960, **60**, 68-75.

Pivik, T., and Foulkes, D. "Dream deprivation": effects on dream content. *Science*, 1966, **153**, 1282-84.

Rechtschaffen, A., Verdone, P., and Wheaton, J. Reports of mental activity during sleep. *Canadian Psychiatric Association Journal*, 1963, **8**, 409-14.

Riley, M. W., and Riley, J. W. A sociological approach to communications research. In Schramm, W. (Ed.) *The process and effects of mass communication.* Urbana: University of Illinois Press, 1954.

Schramm, W., Lyle, J., and Parker, E. B. *Television in the lives of our children.* Stanford, Calif.: Stanford University Press, 1961.

Schultz, D. P. The human subject in psychological research. *Psychological Bulletin*, 1969, **72**, 214-28.

Siegel, A. E. The effects of media violence on social learning. In Baker, R. K., and Ball, S. J. (Eds.) *Mass media and violence.* Washington, D. C.: U. S. Government Printing Office, 1969.

Swanson, E. M., and Foulkes, D. Dream content and the menstrual cycle. *Journal of Nervous and Mental Disease*, 1968, **145**, 358-63.

Walters, R. H. Implications of laboratory studies of aggression for the control and regulation of violence. In Borgotta, E. F. (Ed.) *Social psychology: readings and perspective.* Chicago: Rand McNally, 1969 (orig. 1966).

Walters, R. H., and Llewellyn-Thomas, E. Enhancement of punitiveness by visual and audiovisual displays. *Canadian Journal of Psychology*, 1963, **17**, 244-55.

Warner, W. L., Meeker, M., and Eells, K. *Social class in America.* New York: Harper & Row, 1960.

Weisz, R., and Foulkes, D. Home and laboratory dreams collected under uniform sampling conditions. *Psychophysiology*, 1970, **6**, 588-96.

Whitman, R. M., Pierce, C. M., and Maas, J. Drugs and dreams. In Uhr, L. and Miller, J. G. (Eds.) *Drugs and behavior.* New York: Wiley, 1960.

Wilcoxon, F., and Wilcox, R. A. *Some rapid approximate statistical procedures.* Pearl River, N. Y.: Lederle Laboratories, 1964.

Witkin, H. A., and Lewis, H. B. The relationship of experimentally induced presleep experiences to dreams: a report on method and preliminary findings. *Journal of the American Psychoanalytic Association*, 1965, **13**, 819-49.

Wood, P. Dreaming and social isolation. Unpublished doctoral dissertation, University of North Carolina, 1962.

Yarrow, M. R., Campbell, J. D., and Burton, R. V. Recollections of childhood: a study of the retrospective method. *Monographs of the Society for Research in Child Development*, 1970, **35**, no. 5 (whole no. 138).

Race, Identification, and Television Violence

Cedric Clark

Stanford University

As creators and sustainers of shared perceptions of reality and ways of thinking, the mass media often provide the concepts which are the working tools of the social scientist. When these concepts are, in turn, applied in the investigation of the media, the investigator and his objects of observation become parts of a *closed system*. This operates to the detriment both of the social scientist (who desires a full understanding of the media) and of the media system itself (which needs criticism or "negative feedback" if it is to survive).

It is often necessary that such systems dramatize events if these events are to receive the attention the media need in order to be recognized and maintained as viable social systems. In a real sense, the concepts relating to human behavior are exemplified by the mass media through such dramatization. Nowhere is this perhaps more evident than with respect to the concept of *violence.* The common conception of violence is certainly heavily influenced, if not created, by mass media presentations of examples.

The significant question, then, is: what examples of violence do the media present? How is the concept of violence defined by the mass media, and to what dangers is an investigator exposed in using such a definition?

In most cases, examples of violence presented by the media are those in which the dimensions of *visibility* and *speed or rapidity* are salient. That these two dimensions are the most important in presentations (and hence in definitions) of "violence" is not surprising when we note the importance of drama in the production of mass media messages. The underlying dimensions of *space and time* are also crucial to (Western) man's thought patterns. We tend to emphasize those things which occupy clearly defined points in space (which are visible) and which occur within a certain temporal framework (which have a certain duration).

The conception of violence used in the present paper, by contrast, emphasizes the dialectic elements of the space/time coordinate system; it emphasizes events which are relatively *invisible* and *slow.* Rather than concerning ourselves with *sudden physical death or injury,* which is the typical representation given by the mass media themselves, we will focus on *slow mental disintegration* or *torture.* Only with such a conception are we able to transcend the "blinders" placed upon us by the media and thus enable ourselves to assess the violent consequences of social institutions in terms which do not, logically, exclude the mass media. We must at least entertain the possibility that the mass media do not merely reflect or produce violence, but that they themselves *are* violent.

In brief, our fundamental proposition, or working assumption, is that *the mass media commit violence by virtue of their effects on the Black self-image.*

The Black self-concept is damaged by two parallel sets of forces operating in American society: feelings of *incompetence* are generated by the use of race as a principle of social organization; feelings of *delegitimation* are generated by the communication messages which reflect that social organization.

Three basic propositions are central to our concept of legitimacy:

1. All human systems require maintenance inputs.
2. Legitimation is an important type of maintenance input.
3. Black Americans have been denied legitimation, both from whites and from other Blacks.

Maintenance inputs

Maintenance inputs are those which energize a system and make it possible to process signals (events). A system of mass communication requires two kinds of maintenance inputs: *money* and *legitimation.*

It is obvious that a mass media system cannot exist without adequate sources of financial support. On the face of it, it might be assumed that media systems would accept all potential monetary input, regardless of the source or the nature of the media system itself. This seems not to be the case, however. Whether a media system accepts monetary support as a maintenance input depends on the "state" of the system. For example, if a newscaster's performance is weakened by a shortage of videotape machines, the media system of which he is a part is likely to accept greater maintenance input of the monetary type.

The maintenance input which communicates a sense of *legitimacy* to the system often takes the form of favorable predispositions or attitudes, which extend *credibility* to the system. The proposition being advanced here is that all self-conscious systems (i.e., all human systems) require a nonhostile environment in which to operate effectively. Since the world is not a static place, and since a system's place in it needs constant reaffirmation, extensions of credibility (of legitimacy) are constantly being sought after (McCall and Simmons, 1966).

All media systems are able to store maintenance input, just as they are able to store signals. The extent of this storage is determined by such factors as memory capacity, physical facilities, etc. With respect to monetary inputs, for example, a mass media system may sell some stock, land, or other facilities and thereby receive additional "operating capital" needed for signal processing. It is thereby capable of generating its own maintenance input. Obviously, however, the supply of this is not unlimited and must therefore be periodically supplemented by additional sources.

Legitimizing maintenance inputs, however, cannot be internally generated. That a mass media system thinks well of itself, has favorable predispositions toward itself, is, at some minimal level, a necessary condition for its ongoing operation, but this minimal level cannot be significantly increased through internal processes. A system can be legitimized only by other systems; it is this exchange of legitimation which leads to the creation of systems.

When it reacts to specific events or specific people in its environment, a media system in a very real sense communicates legitimacy to them by recognizing them. But at the same time, the system is able to gather maintenance inputs by the process. The publication or televised presentation of an interview or a story may spur additional consumption of the message product, thereby granting the system monetary maintenance input through increased advertising revenue. The publication or broadcast may also cause readers or viewers to have more favorable feelings toward the media system.

In other words, certain people and events are *legitimation resources*, and the media act toward them in ways aimed at insuring that this supply of legitimation does not fall below some minimal level. These sources of legitimation can be termed "relevant receivers" (Clark, 1970; Easton, 1965b)—those whom the system, from past experience, considers important sources of maintenance input.

The communication of legitimacy is characterized by two distinct activities: recognition and respect. These activities describe the actions taken by one system in relation to another.

Recognition: the first stage of legitimation

To recognize something or somebody means that I pay attention to it; I take it into account and acknowledge its existence. Recognition, in this sense, is related to but distinct from *awareness*. I can be aware of something, yet refuse to recognize it. The element of *choice* is thus implicit in the act of recognition (Clark, 1971b). Suppose, for example, an acquaintance of mine enters a restaurant where I'm sitting. I may be fully aware of his entry, yet I may choose to ignore him. My acquaintance may, of course, *impose* his presence on me and, under such conditions, I may be forced to accord him recognition (but not necessarily respect). When I make this act of recognition, I have provided the first (and necessary) condition for his legitimation.

This example illustrates that recognition can operate as a form of social (and political) control. By not recognizing somebody, we can devaluate his importance and hence keep his behavior under control (at least to the extent that the other person desires this recognition from us and can receive it only from us).

The opposite of recognition is, then, nonrecognition. If the English language had a satisfactory noun form of the verb "to ignore" ("ignorance" has, of course, misleading connotations), such a term would be quite useful in describing what is meant here. It would connote awareness of, but nonrecognition of.

The psychological importance of recognition. William James notes:

> A man's social "me" is the *recognition* which he gets from his mates. . .no more fiendish punishment could be devised. . .than that one should be turned loose in society and remain absolutely unnoticed by all the members thereof (1961, p. 469; emphasis added).

James goes on to emphasize that it is only through recognition by others that a person's sense of "self" is able to develop—a point Mead

(1934) also stresses. The interaction with others enables one to know himself as an object as well as an actor. If other people did not take one into account (i. e., recognize one), it is doubtful that a concept of the self would develop. A person would not be able to differentiate himself from his environment. Such differentiation is important if man is to avoid the "ontological insecurity" described by Laing (1960).

The form of recognition. How does one know if he is being legitimated (recognized)? Evidence of recognition takes two basic forms, depending upon whether the communication process is interpersonal (face-to-face) or impersonal (mass). In the first instance, recognition occurs in the form of verbal address—e. g., "John, how are you?" Through such direct or immediate communication, the existence of John is recognized, and the necessary condition for legitimation has occurred. (It is not necessary, however, for the personal noun ["John"] to be included in the address, since circumstances usually indicate quite clearly that it is John who is in fact being recognized. If I pass John walking down the street and say "hello," or "what's happening, man?", I have recognized him in the form of verbal address.) It is possible for me not to recognize him (though, again, I may be aware of him). Thus, the statement, "I saw him, but I refused to recognize him" makes sense. It means that, while I was aware of him, I *chose* not to let him know that I saw him. If enough people chose to ignore him, he would be unable to build and maintain a conception of himself.

In a complex society such as the United States, the most prevalent form of recognition occurs not through interpersonal or verbal address but through impersonal symbolic (nonverbal) representation. This form is transmitted through the channels of the mass media of communication and manifests itself in a variety of ways; one of the most important is television content.

In a completely homogeneous society—a society in which there were no major physical differences between people—everyone could be recognized (legitimated) through pictorial or symbolic representation, because a single pictorial representation would resemble all concerned. It is obvious, however, that in complex, multiracial societies there are physical differences between people; therefore, it is impossible to recognize all groups through a uniform nonverbal representation. Hence, some people (systems) tend to be nonrecognized (or delegitimated). The consequences of this in American society—where whites provide most of the maintenance inputs (money and legitimation) to the mass media—is the general nonrecognition or delegitimation of Black Americans.

The determinants of recognition. What determines whether one system (A) receives recognition from another system (B)? While there are perhaps numerous determinants, the most important appear to be the *uniqueness* or informational value of system B's actions, and the perceived *relevance* of B's actions to the goals and values of A.

The notion of uniqueness is related to the distinction made earlier between awareness and recognition; to be aware of something or to have knowledge of something does not necessarily mean that one recognizes it in the sense of "paying attention to it" or "taking it into account." It is assumed that there is some minimal level of awareness of system B's actions on the part of A, the observing system. Given this level of awareness, what may determine whether A chooses to pay attention to B is the extent to which these actions are unique or unusual.

The uniqueness of an action or event is no guarantee that it will be taken into account, however. The action must also be *relevant* if an observing system is to pay attention to it. A communication from one individual *addressed to* another individual is clearly relevant. Such address singles out or differentiates an individual. Thus, in a group situation, I will pay attention to somebody if he calls (addresses) me. If he does not —if he speaks to the "group in general"—it is easy to ignore him.

When the form of recognition is mediated (impersonal), rather than immediate (face-to-face), relevance is determined less by address than by the perceived effect of the actor's behavior on the goals or values of the observer. The fundamental goal of all individuals is self-maintenance; thus, actions viewed as affecting an individual's self-concept will have a high probability of being recognized. This enables the individual to defend himself from the danger presented. An action has to be goal-related if it occurs in impersonal form, because the communication receiver (the observer) has greater "degrees of freedom" to ignore it. I can, for example, throw away (nonrecognize) a piece of "junk" mail. I will pay attention to it only if it is clearly related to some personal goal or value.

The reason most American mass media systems do not pay attention to Blacks is that Blacks are the "irrelevant receivers"; they do not provide the maintenance inputs needed by the systems. Only when the activities are perceived as directly related to the goals or interests of media systems (as, for example, the riots of the 1960s were) are such actions generally recognized.

Respect: the second stage of legitimation

Recognition is a necessary condition for legitimation to occur, but not a sufficient one. Without recognition, legitimation is impossible, but it alone is no guarantee that legitimation will occur. For example, a particular individual may recognize or take into account another individual, but in a way which may communicate disrespect and effectively *delegitimize* him. Particular kinds of message content may also communicate disrespect.

Message content which communicates respect is of three types: (1) messages which share the *definition* of an actor's behavior; (2) messages which share the *assessment* of an actor's behavior; and (3) messages which share the *accountability* of an actor's behavior. To facilitate exposition, it is worth noting here that an individual who is "respected" (and hence legitimated) by an observer or observing system is one (1) whose own definitions are employed by the observer in the interpretation of the actor's behavior; (2) whose own assessments are employed by the observer; and (3) whose own behavior, if favorable, is generally accounted for by attributing it to the actor's dispositions and, if unfavorable, is accounted for by attributing it to the individual's environment.

These three types of message arrangements, together with the first stage of legitimation (recognition), are diagrammed in Figure 1.

STAGE 1 (RECOGNITION)	STAGE 2 (RESPECT)
An individual identifies the existence of another by:	An individual identifies with another by:
1. Paying attention to him; by taking him into account	1. Sharing his definitions of his behavior
	2. Sharing his assessment of his behavior
	3. Sharing his explanations for his behavior

Figure 1: The process of psychological legitimation

Respect and behavior definition. The way a person's behavior is defined by another has important consequences for how both act in relationship to each other. When person A exhibits behavior X in front of person B, it is unlikely that A and B will perceive (and hence define) the behavior the same way, because they bring different experiences and sets of expectations to the situation. There is thus always some variation in the ways any two individuals perceive and define an action. However, such variation can be minimal or maximal, and these limits often communicate "respect" and "disrespect," respectively.

Respect is communicated by an observer when his definition of the actor's behavior shows little variance with that held by the actor himself. Lack of agreement in definitions is often a sufficient condition for the communication of delegitimation; it is rare that a person is able to legitimate another if that person and he do not define reality in similar ways. This is not to suggest that an *identity* of definitions is common or, perhaps, always desirable. There will always be some "error variance" which serves to protect the actor's own "core" identity.

What is suggested, then, is that not only is it important for a person to have his behavior recognized by an external observer, but it is also important that his actions be defined by others in a way consistent with his own—at least to some minimal extent.

Not only do shared definitions enable A to predict B's actions; they also inform A about how to behave in the presence of B (Newcomb, 1953). Shared definitions are, then, both *informing* and *maintaining*. Their importance is summed up in the recognition that psychological (and other) systems which lack either information (an informing function) or energy (a maintaining function) will disintegrate (Easton, 1965a; Berrien, 1968). Computer systems, for example, require both signal or data (information) and maintenance (electrical energy) for effective functioning. Similarly, man requires both information and maintenance to deal with his environment.

Like the process of recognition, the message elements which constitute the definition of an act can exist in either interpersonal (face-to-face) or impersonal (mass media) communication forms.

In interpersonal communication forms, behavioral definition occurs when the observer (the communication source) employs linguistic concepts to encode some aspect of another person's behavior. Since encoding itself involves the more or less conscious choice of labels or concepts for the classification of behavior, the process of definition is inherent in it.

In impersonal or mass communication forms, behavioral definition occurs when the observer (in this case, a mass media system) encodes and exemplifies some aspect of another person's (or group's) behavior. Rarely is a person himself presented in the mass media; he is usually represented by somebody else, and these representations are usually fictional and dramatic. Such dramatizations carry with them explicit definitions of actions. A good example is the actions of American Indians as they have been defined by mass media. A battle waged and won by Indians is frequently defined as a "massacre," while one won by non-Indians is usually defined as a "victory."

The essence of defining, then, is the use of concepts in the process of classifying social reality. Because concepts (and, hence, definitions of behavior) originate in human experience, they are frequently peculiar to a given set of experiences. To the extent that experiences differ substantially (like, for example, those of Black and white Americans), the concepts derived from one set may be quite different from those derived from the other set, even when the behavior observed is "identical." What we perceive is culturally determined. One observer looking at the family structure of Black Americans may see weakness, while another observer, from a different cultural perspective, may see strength (Billingsley, 1968).

Simply because a person is born into one culture which has concepts different from another, this does not mean that he can never learn the conceptions and definitions of that other culture. This is achieved via second language learning. It is hence clear why learning another person's language is often said to be one of the highest tributes (respect) one can give. It represents a *deliberate* attempt to learn how another group of people perceive reality.

Applying this notion directly to the nature of Black/white social interactions, it will be recognized that most white Americans do not, for various reasons, learn "Black English" as well as Blacks learn "White English" (Taylor, 1970). Such a situation often leads to a process of delegitimation of Black behavior on the part of whites.

Respect and behavioral assessment. Concomitant with the process of defining an action is the process of evaluating it or assessing it. That is, to the extent that "meaning" is derived through definition, a strong evaluative component is likely to be part of this meaning (Osgood et al., 1957).

Much of man's behavior involves the evaluation and assessment of others. The assessment need not be manifest or, for that matter, conscious. But the tendency to evaluate seems part of the very process of attaching meaning to an object or activity. (The term "assessment" is perhaps preferable to "evaluation" inasmuch as the former tends to encompass a much wider range of judgments than simply "good or bad". Other dimensions, while still strongly evaluative, are relevant in different contexts—e. g., "successful-unsuccessful," "competent-incompetent," "relevant-irrelevant." Which dimension is most appropriate depends on the nature of the particular behavior manifested.)

People not only want to be recognized as existing and to have their actions defined in cognitively similar ways; they also want their actions to be evaluated or assessed. People want to be supported in their own classifications and evaluations of reality, and such support cannot be internally generated because an act of *comparison* is required (Festinger, 1954). Such comparison can only be acquired by interacting or communicating with others or, at least, by having them communicate with you.

In addition, assessment functions as "feedback " (positive and negative) for psychological systems. Such feedback is necessary for the successful functioning of all behavioral systems. Without adequate feedback, an individual will never know whether or not he is successful in the pursuit of his goals.

Like recognition, assessment can occur in both interpersonal (face-to-face) and impersonal forms. In the former case, an individual may simply say to another, "Your behavior is good."

Assessment may, however, take forms other than the obvious. If I wanted to negatively evaluate a person's actions, I might, for example, speak in a sarcastic manner of his achievements. Ridicule is, in fact, a very common method of communicating delegitimation and one frequently used in television. This suggests, again, that many assessments are not direct, but are mediated through impersonal or mass communication channels: on television, in books, in magazines, etc. In addition to (or instead of) labeling a particular action as "good" or "bad," the mass media may employ more implicit devices. Many messages in the mass media are presented to us through prestructured forms, such as drama or athletics. Through such forms, people or groups can be "cast" in the form of "heroes," "villians," or "fools" (Klapp, 1962).

In a television presentation, an individual's actions may be placed sequentially between actions which are themselves labeled as "bad." Through the process of association, the viewer learns that the action in question is "bad." By placing a report of a civil rights demonstration between reports of clearly labeled criminal activities, the event tends to be negatively assessed.

Stylistic elements associated with pictorial representation are also important in communicating respect. For example, presenting a picture of a man smiling tends to favorably assess the behavior associated with him; presenting him frowning leads to a negative assessment of his behavior.

An example of delegitimation through pictorially produced assessment is the cartoon product known as a caricature. The exaggeration of physical features tends to influence one's judgment of the subject's behavior—in ways which are usually unfavorable. The dimension applied to the assessment of actions associated with a person thus caricatured is "serious/nonserious." What one learns through such portrayals is that the individual's behavior is "not to be taken seriously." Not to be taken seriously is a form of delegitimation common to many Black Americans. The television images of Blacks presented in such programs as *Stephin Fetchit, Tarzan*, and *Amos 'n Andy* may be cases in point.

The way the actions of a particular individual are assessed depends on: (1) the type of assessment *dimensions* possessed by the observer; (2) the *uniqueness* or visibility of the actions; and (3) the *relevance* of the actions to the observer.

Assessment requires some dimensions or criteria used in the process of evaluation. Such dimensions often tend to be specific to the action displayed; the dimensions we use to judge how well a person plays football are different from those we use to assess how well he drives a car. While a skill is involved in both activities, it is likely that different dimensions associated with this "skill" will be employed. For example, we may talk of "agility" with respect to football playing and "care" with respect to driving a car. It is important to note, however, that, inasmuch as dimensions are concepts, they are learned just as other concepts are learned. Thus, individuals who are observers of a culture distinct from their own may lack the dimensions relevant for the assessment of the behavior manifested in that culture.

The uniqueness of the actions taken by an actor also affect the nature of behavioral assessment. The behavior must fall within certain culturally known boundaries if it is to be assessed at all by an observer. If the behavior is too unusual, it will be impossible for the observer to evaluate it. Many examples of creativity, for example, remain unappreciated because the requisite dimensions for assessment are unavailable to observers.

Thus the ultimate determinant of behavioral assessment is the cultural values of the observing individual. Values are the standards which tell us what is relevant and why things are worth pursuing. They heighten our attention to certain events and serve as standards of judgment. They are thus very important in behavioral assessment.

Respect and behavior accountability. The third component of respect, *accountability*, is closely associated with the process of assessment or evaluation. Accountability involves the attribution of blame or credit for the occurrence of an individual's actions. The tendency for man to search for causes of events—both social and nonsocial—is pervasive and may in fact be considered a fundamental psychological process (Heider, 1957).

The tendency to attribute causality has recently been discussed in considerable detail by Kelley (1967), who notes that, in general, either an observer holds the individual accountable for his actions and attributes causality to his dispositions, or he holds a person's environment accountable and attributes the cause of his actions to external forces. Attribution can thus be of two basic types: internal or external (Rotter, 1962).

People tend to use both external and internal cues to attribute causality, but one orientation often predominates with respect to a particular class of events. In addition to recognizing, defining, and assessing, man also seeks structure or predictability in his environment. He can do this by explaining (attributing cause to) specific events in that environment.

External attribution is frequently accomplished by placing the action in a temporal or a spatial message sequence such that it is preceded by some event or set of events. The antecedent action is thus interpreted as a cause. For example: "A banana peel lay on the sidewalk." "A man fell on the sidewalk." The sequence of two sentences carries a meaning quite different from that which would be conveyed if the second statement appeared alone. In short, the appearance of the first enables one to "explain" (attribute causality to) the second.

Internal attributions of causality, then, can be suggested by the absence of an external attribution or "explanation." They can also be communicated by applying invariant trait descriptions to the acting individual. Stereotypes, for example, serve to "explain" the behavior of an individual in terms of his internal state. Thus, a person is considered "clumsy" if he falls down.

To fully understand the determinants of the specific types of accountability (internal and external), consider the kinds of attributions a person normally makes for his own behavior. Whatever the particular attributions he makes—whether internal or external, they are likely to be consistent with his basic goal: to maintain a satisfactory self-concept. Thus internal accountability will probably occur when a person likes his actions or considers them "successful," and external accountability will occur when a person dislikes his actions or considers them "unsuccessful."

For example, suppose I fall down while walking. An observer might witness this event and in some manner express recognition of it; he might agree with my *definition* of the act as a "fall"; he might further agree with me in assessing the fall as an "unfortunate" event. When it comes to *accountability*, however, I, as an actor, may attribute causality in a different way from the observer. This may be because I am aware of more or different cues than the observer (who may be located at some distance). However, it is more likely that, in order to maintain "face," I will attribute such incompetent behavior, not to my own internal failings, but to some external force (a banana peel or a "trip" by someone).

The relevant proposition is: the more negatively assessed the behavior, the less likely a person is to accept accountability (blame) for it; conversely, the more positively assessed the behavior, the less likely a person is to attribute causality to (credit) external forces. In other words, we like to take credit for the good things we do and to blame others for the bad things we do. I slipped because the sidewalk contained a banana peel, not because I was "clumsy." By making such attributions, I am able to take account of my own behavior in a way that maintains my self-concept as a nonclumsy individual.

ASSESSMENT OF ACTOR'S BEHAVIOR

		Competent (Positive)	Incompetent (Negative)
SOURCES OF CAUSALITY (ACCOUNTABILITY)	Internal (Personal)	RESPECT (High Identification)	DISRESPECT
	External (Environmental)	DISRESPECT	RESPECT (High Identification)

Figure 2: Conditions under which second stage of legitimation is fulfilled (assuming shared "definitions")

In the process of legitimation, however, the important question is not individual attributions of one's own behavior, but the attributions one gives to others. Such other attributions are fundamental to the communication of legitimacy.

The proposition presented here is that to the extent that individual A respects individual B, individual A will account for B's behavior in the same way as B would account for his own behavior. Thus, assuming a basic need for self-confirmation, good or positively assessed actions will be attributed to the internal characteristics of another, while bad or negatively assessed actions will be attributed to environmental forces. In other words, if we have respect for a person, we will blame his failures on things over which he has no control and credit his successes to things intricately associated with him. This defines the extent to which we can *identify with a person.* To the extent that we do, we communicate respect or legitimacy to him. Figure 2 illustrates the type of accountability actions possible.

THE MEDIA, BLACK CONSCIOUSNESS, AND LEGITIMATION

"Black consciousness" or "Black awareness" can be viewed in a cognitive state which actively reconsiders (a) the sources, (b) the form, and (c) the determinants of personal and cultural legitimation.

Legitimacy and black consciousness: recognition factors

While the empirical evidence is sparse, one might argue that the 1970s will be characterized by a shift in the *sources* of Black legitimation from white to Black. Greater numbers of Black Americans will actively seek recognition from other Blacks and, in turn, will communicate recognition to them. The dominant interactants in this new communication system are those who possess some degree of "Black consciousness."

As indicated earlier, recognition takes both personal (face-to-face) and impersonal (mass mediated) forms. Increases in Blacks communicating with other Blacks are evident, particularly on college campuses and in educational institutions in general. Calls for varieties of "Black separatism" reflect this increasing willingness for Blacks to pay attention to other Blacks (which some whites interpret as an unwillingness to pay attention to them). While hardly extensive, there also appears to be greater recognition of Blacks by Blacks in the mass media. Subscriptions to Black newspapers and magazines, for example, appear to be increasing: *Muhammad Speaks,* the official newspaper of the Nation of Islam, is now the largest-selling Black newspaper.

The determinants of such recognition have been discussed in terms of the *uniqueness* and *relevance* of an individual's actions. In the context of the present Black movement, actions by other Blacks seem to have higher informational value and relevance than they did in previous decades. In other words, what other Blacks are doing, saying, and thinking now "matters" more or "makes more of a difference" to Blacks. The *necessary* condition for Black legitimation by other Blacks is thereby achieved.

Legitimation and black consciousness: respect factors

Recognition is a necessary, but not sufficient, condition for the process of legitimation to occur. What is also required is the communication of respect: a sharing of definitions, behavior assessments, and causal attributions (accountability). In the context of today's Black struggles, the communication of respect to other Blacks appears to be a more difficult task than that of recognition. Nevertheless, there is currently widespread emphasis on the sharing of "Black" definitions, assessments, and attributions. Since all of these have their ultimate origin in human experience, there are more determined efforts to identify, articulate, and communicate the distinctive aspects of this experience.

With respect to behavior *definition*, perhaps the most dramatic form of concentrated efforts in this direction is the change in self-definition: "Black" Americans define themselves differently than "Negro" Americans. This redefinition of what or who one is is a kind of religious conversion; one is, in effect, "reborn" (cf. Thomas, 1971). Such rebirth or redefinition leads many Blacks to communicate respect to other Blacks by defining (or redefining) their behavior in terms of "Black" (as opposed to "white" or "Negro") conceptions of reality. The terms "Black," "Nigger," "African," "tribal," etc. take on new meanings and play an important part in both informing and maintaining Black psychological systems—in providing both signals and maintenance.

Part of the new Black consciousness involves the *reassessment* as well as the redefinition of Black behavior. Actions which were at one time assessed as "bad" are currently under reassessment; either they are being defined as "good," or judgment is being suspended. The behaviors which were previously denied or repressed are now openly displayed as evidence of a "new breed" of man. Having "nappy hair," "thick lips," or other visible evidences of being of African descent are no longer viewed as "stigmas" (Goffman, 1963); they are assessed as positive aspects, if not the essential ingredients, of "Blackness." There is now often a direct (negative) correlation between what whites and Blacks assess as positive or "good" behavior. Such differential definitions of reality lead to unpredictability in social relations and, thereby, make changes in the social system possible (if not inevitable).

But, because a level of negative self-evaluation is still present in many Blacks, there is often an inability to identify with Blacks. Society (largely through the mass media) presents mostly whites as models for identification. To the extent that the "perception of common qualities" is an important determinant of identification, the presentation of only white models makes identification difficult if not impossible for Blacks.

Thus Blacks (as well as whites) tend to account for the success of other Blacks in terms of an "enriched environment" (external attributions) rather than "innate genius" (internal attributions). Conversely, Black failures are often still accounted for by internal causality (commonly referred to as "stereotypes"), rather than in terms of external conditions (e.g., "police harassment"). This problem has been recognized by many Blacks and is related to the development of "Black studies" programs designed to provide the environmental and nonenvironmental information necessary for generating patterns of accountability capable of eliciting greater Black respect for Black behavior. (Robinson et al., 1969)

Black consciousness and the mass media

The preceding remarks can be epitomized with the statement that the chief distinctive element of "conscious" Blacks is a strong determination to *identify with other Blacks*.This raises some interesting questions about how such people view the relationship of the (white) media system to this objective. Are the media viewed as helpful or harmful? relevant or irrelevant? How are the Black characters presented on television and how do such Black viewers perceive these?

Some of these questions are addressed in the research reported in the present paper. The studies represent an attempt to test specific propositions and assumptions which have thus far been taken for granted. The various concepts discussed here in a theoretical vein—identification, Black consciousness, legitimacy, etc.—are made operational, and specific hypotheses are tested. These studies will be more readily interpretable if one crucial thought (or prejudice) is kept in mind: *mass communication in general, and television in particular, have their most violent effect in terms of how they have defined, assessed, and explained Black behavior.*

THE CONCEPT AND MEASUREMENT OF BLACK CONSCIOUSNESS

We conceived of Black consciousness as having two basic components: (1) the extent to which one feels that he has been "programmed" —i. e., the extent to which he recognizes that his role as a Black has an ascriptive base; and (2) the extent to which he attempts to redefine this role. Blacks themselves popularly refer to this process as "getting oneself together."

Within this two-component conceptual framework, two other subcomponents were postulated and, within these, a series of "change-related" beliefs were identified. All of these are summarized in Table 1.

While the conceptual framework has some definite features unique to Black Americans, it should be noted that our theoretical rationale logically includes such phenomena as "women's consciousness," "youth

Table 1: Theorized components of black consciousness

GENERAL COMPONENTS:

1. Recognition of social role learning ("programming")
2. Redefinition of ascribed role ("reprogramming")

SPECIFIC COMPONENTS:

1. Concern with physical appearance
2. Concern with behavioral characteristics

COGNITIVE CHANGE PROCESSES:

1. Denial to recognition of physical and behavioral patterns
2. Liability to asset of physical and behavioral patterns
3. Shame to pride of physical and behavioral patterns

consciousness," etc. In other words, the concept of "consciousness" presented here springs from the types of social awareness associated with specific points in American social space. Society teaches us the preferred modes of behavior which characterize whites *vis-a-vis* Blacks, men *vis-a-vis* women, and old people *vis-a-vis* young people. The extent to which one is aware of such teachings, and does not merely accept them as "natural" or "given," is a major component of social consciousness.

Being aware of such societal teachings does not, however, in and of itself mean that one is dissatisfied with them or that one attempts to alter their effect on one's own identity. The second component, role redefinition, occurs when one is, at some level, alienated from the social identity one has been given.

We have exemplified both of these general components by citing specific attitudinal items designed to tap specific elements and the cognitive change processes associated with them. These latter change processes include: (1) a change from denial to recognition of physical or behavioral differences; (2) a change from viewing one's physical and behavioral attributes as liabilities to viewing them as assets; and (3) a change from viewing them with shame to viewing them with pride. In general, the items refer to a change from self-rejection to self-acceptance.

The specific items concerning these three processes are shown in Tables 2-4.

The concept of "ascribed role consciousness" as applied to Blacks was conceptualized as containing four additional components which serve to distinguish it from the consciousness associated with age and sex as organizing principles of society. In brief, these additional components take into consideration the temporal aspects of Black social behavior, particularly in relationship to what is commonly referred to as the "Black movement." Specifically, these components include: (1) the

Table 2: Pilot test items referring to denial-to-recognition change processes with respect to role learning and role definition

Role learning	Role redefinition
A. Physical characteristics	A. Physical characteristics
25. The reason why some members of my race want to look white is because of what they have seen on the mass media.	1. The color of skin of most members of my race is quite different from that of the white race.
27. The reason why some members want to look white is because of what they have learned from their parents.	3. The shape of the nose of the average member of my race is quite different from that of whites.
29. The reason why some members of my race want to look Black is because of what they have learned in most schools.	5. The lips of most members of my race are no different from those of whites.
	7. The hair texture of members of my race is basically the same as that of whites.
B. Behavioral characteristics	B. Behavioral characteristics
26. The mass media fail to recognize differences in behavior between whites and members of my race.	2. The behavior of most members of my race is quite different from that of whites.
28. Most schools do not recognize differences in behavior between whites and members of my race.	4. Most members of my race dress quite differently than do whites.
30. The parents of most members of my race emphasize differences in behavior between themselves and whites.	6. The way members of my race eat is no different from that of whites.
	8. Most members of my race do not walk any differently than whites.

perception of the movement's *goals;* (2) the *means* of achieving them; (3) the *obstacles* in the way of achievement; and (4) the *concepts* reflecting the space/time cultural dynamics. Examples for each of these are presented in Table 5.

Preliminary studies

Subjects. The various items listed in Tables 2-5 were incorporated into a general questionnaire. Respondents were asked to indicate the extent of their agreement along a five-point Likert-type scale. Some of the items were reversed so as to preclude the formation of a response set. The phrasing "members of my race" was employed because at the time of the study (1970) the various terms used to label Black Americans were themselves an important variable; the variable was in fact incorporated into our instrument (item 48).

The Black consciousness scale was administered to a group of Black Stanford undergraduates (N = 17) and Black high school students from

Table 3: Pilot test items referring to liability-to-asset change processes with respect to role learning and role redefinition

Role learning	Role redefinition
A. Physical characteristics	A. Physical characteristics
31. The reason some members of my race want to look white is because of what they have been taught by their parents.	9. A dark skin is more beneficial than a white one.
33. The reason some members of my race consider their skin color a handicap is because of what they have seen in the mass media.	11. A wide, flat nose is more beneficial than a thin, narrow one.
35. The reason some members of my race consider their skin color a handicap is because of what they have been taught in school.	13. Soul food is better for you than other types of food.
	15. There are no definite benefits associated with thick lips.
B. Behavioral characteristics	B. Behavioral characteristics
32. The reason some members of my race feel inadequate in their behavior is because of what they have learned from their parents.	10. High athletic ability has been of great benefit to members of my race.
34. The reason some members of my race feel inadequate in their behavior is because of what they have seen in the mass media.	12. The type of hair members of my race have is no more beneficial than that of whites.
36. Most schools have taught members of my race to feel inadequate in their behavior.	14. Soul music is no better than other types.
	16. Dashikis and other African-type clothes are less comfortable than regular clothes.

an all-Black school (N = 11). The samples were almost evenly divided with respect to sex, with one more male than female in each group. Most students were in their third year at the two institutions. A Black female experimenter administered the questionnaire to the high school sample and a Black male to the college sample; in both cases, the administrator was the teacher in charge of the class.

Data analysis. In order to examine the structure of relationships among the various items (which was our only objective in this study), the various responses were intercorrelated. While separate analysis was done on the two samples, the structuring of responses was virtually the same; there were some differences in other respects, however, and these will be indicated. The following Tables 6-9 thus represent the Pearson Product-Moment correlation within each of the conceptualized components, with degrees of freedom at 27 (necessitating a correlation value of .31 for significant chance deviation at the .05 level).

Role redefinition: behavioral characteristics. Table 6-A shows that only two of the four variables are significantly related. These represent

Table 4: Pilot test items referring to shame-to-pride change processes with respect to role learning and role redefinition

Role learning	Role redefinition
A. Physical characteristics	A. Physical characteristics
37. The reason why some members of my race feel ashamed of their skin color is because of what they have learned in the mass media.	17. A broad, flat nose is not as attractive as a narrow one.
39. The reason why some members of my race feel ashamed of their skin color is because of what they have learned in the mass media.	19. Members of my race should be ashamed to straighten their hair.
41. Most schools teach members of my race to be proud of their skin color.	21. Members of my race should not be too proud of the color of their skin.
	23. Members of my race ought to be proud of their thin lips.
B. Behavioral characteristics	B. Behavioral characteristics
38. Most schools have done a lot to destroy pride in the behavior of members of my race.	18. Members of my race should never be ashamed of the way some members behave.
40. The mass media have done a lot to make members of my race ashamed of their behavior.	20. Members of my race should be proud to do such dances as the "dog."
42. The parents of most members of my race have instilled pride in their children's behavior.	22. Members of my race ought to be ashamed of their music.
	24. Members of my race ought to be ashamed of the food they eat.

responses to the questions regarding racial differences in eating and walking.

Table 6-B shows two independent clusters with responses to the questions concerning the benefits of athletic ability and "soul food" having a significant positive correlation. The "soul food" item revealed a negative correlation with a negatively worded item concerning the benefits of African-type clothing; this, of course, was expected.

Table 6-C, concerning the cognitive shift from a "shame to pride" orientation, revealed no significant intercorrelations with the other items.

Role redefinition: physical characteristics. Table 7-A shows a significant relationship between items concerning recognition of "hair texture" and "lips." Interesting, however, the items concerning skin color did not relate to any of the others in this cluster.

Table 7-B shows two distinct clusters of related variables: one set relates the items concerning "skin color" and "shape of nose," and the other set relates "hair texture" and "shape of lips."

Table 7-C shows only one pair of variables significantly related. These concern "nose attractiveness" and "hair straightening." The negative correlation is consistent with theoretical expectations.

Table 5: Items relating to additional components of black consciousness

Goals:

46. The goal for people of my race should be total integration.
50. The goal of members of my race should be essentially the same as with other ethnic groups.

Means:

47. Members of my race should try to achieve their goals by any means necessary.
51. Increased education is the best way for members of my race to achieve their goals.

Obstacles:

43. One of the things standing in the way of my race's achieving its goals is the emphasis on integration.
44. One of the things standing in the way of my race's achieving its goals is the adoption of Christian values.
45. One of the things standing in the way of my race's achieving its goals is lack of knowledge regarding the way political processes work.

Concepts:

48. The terms "Black," "Negro," "Afro-American," and "Colored" all mean the same.
49. All whites are basically the same when it comes to reacting to members of my race.

Table 6: Intercorrelations among role redefinition black consciousness items: behavioral characteristics

A. Denial to Recognition

	2	4	6	8
2 Behavior				
4 Dress	23			
6 Eating	–13	–22		
8 Walking	–03	–09	40	

B. Liability to Asset

	10	12	13	16
10 Athletics				
12 Hair	15			
13 Food	56	05		
16 Clothes	13	02	–37	

C. Shame to Pride

	18	20	22	24
18 Behavior				
20 Dance	18			
22 Music	07	–14		
24 Food	18	–24	28	

Table 7: Intercorrelations among role redefinition black consciousness items: physical characteristics

A. Denial to Recognition

	1	3	5	7
1 Skin				
3 Nose	−16			
5 Lips	05	11		
7 Hair	03	10	35	

B. Liability to Asset

	9	11	12	15
9 Skin				
11 Nose	38			
12 Hair	−24	11		
15 Lips	−27	−09	39	

C. Shame to Pride

	17	19	21	23
17 Nose				
19 Hair	−43			
21 Skin	06	−17		
23 Lips	20	−13	−20	

Table 8: Intercorrelations among role learning black consciousness items: behavioral characteristics

A. Denial to Recognition

	26	28	30
26 Media			
28 Schools	38		
30 Parents	12	12	

B. Liability to Asset

	32	34	36
32 Parents			
34 Media	31		
36 Schools	22	79	

C. Shame to Pride

	38	40	42
38 Schools			
40 Media	50		
42 Parents	10	13	

Role learning: behavioral characteristics. Turning now to the second major component of Black consciousness, Table 8-A shows a significant relationship between awareness of the mass media's denial of behavioral differences and that experienced in educational institutions.

There is a relatively strong cluster of variables related to the role of the media and schools with regard to teaching Blacks to feel that their behavior is a liability (Table 8-B).

Table 8-C again shows a high correlation between the social learning gleaned from the schools and that from the mass media. It is interesting to note that the role of parents in this socialization process shows no relationship with the other dominant socializing agencies.

Table 9: Intercorrelations among role learning black consciousness items: physical characteristics

A. Denial to Recognition

	25	27	29
25 Media			
27 Parents	61		
29 Schools	13	−25	

B. Liability to Asset

	31	33	35
31 Parents			
33 Media	15		
35 Schools	39	39	

C. Shame to Pride

	37	39	41
37 Parents			
39 Media	66		
41 Schools	04	−16	

Role learning: physical characteristics. Data concerning the other aspect of role learning—awareness of the social meaning attached to one's physical appearance—are presented in Table 7. In all three change processes, the mass media are significantly related to at least one other source. Tables 9-A and 9-C show that the parents and the media are viewed as sharing some commonality with regard to social teaching; in Table 9-B, the parents are seen as similar to the schools, and the schools as similar to the media.

Goals, means, obstacles, and concepts. Since the purpose of the data discussed above was exploratory, no further statistical analyses were made of the structure of interrelationships among the various variables. Examination was made, of course, of such relationships; but this was done chiefly for the purposes of developing a more parsimonious and refined scale. Thus, several criteria, not all of them mathematical (statistical), entered into the final selection of questionnaire items.

Before discussing this in more detail, however, it is worth indicating some of the mean values associated with the various items in the questionnaire. These are presented primarily to show some of the rationale involved in the construction of our final scale.

The five items with the highest and lowest mean values are listed in Tables 10 and 11. In the left-hand margin are symbols indicating whether the particular item belongs to the hypothesized clusters of Role learning (L) or Role redefinition (R); whether it emphasized physical attributes (P) or behavioral characteristics (B), and whether it dealt with change process 1 (Denial to Recognition), change process 2 (Liability to Asset), or change process 3 (Shame to Pride).

Table 10: Black consciousness items showing greatest agreement (N = 28)

Cluster	No.	Item	Mn.	S.D.
RP-1	1	The color of skin of most members of my race is quite different from that of the white race.	4.67	0.94
RP-1	3	The shape of the nose of the average member of my race is quite different from that of whites.	4.54	0.79
LP-1	25	The reason why some members of my race want to look white is because of what they have seen on the mass media.	4.32	1.37
LP-3	40	The mass media have done a lot to make members of my race ashamed of their behavior.	4.32	1.02
LP-3	38	Most schools have done a lot to destroy pride in the behavior of members of my race.	4.28	1.24

The various items relating to the more contemporary aspects of the Black movement showed very little intercorrelation in the hypothesized pattern. However, all of them were in the expected (positive or negative) direction. Some are discussed more fully below.

Three general characteristics seem to emerge from examination of the ten items listed in Tables 10 and 11. They concern issues involving role *redefinition* more than role learning. They emphasize a concern with *physical characteristics* as opposed to behavioral ones. They concern the change processes of "Denial to Recognition" and "Shame to Pride" rather than the "Liability to Asset" process. These findings played an important role in determining the selection and reformulation of items appearing in the refined scale discussed in the next section.

An examination of mean values for each of the items showed that only one item significantly differentiated the college sample from the high

Table 11: Black consciousness items showing greatest disagreement (N = 28)

Cluster	No.	Item	Mn.	S.D.
RP-3	21	Members of my race should not be too proud of the color of their skin.	1.32	0.98
RB-3	22	Members of my race ought to be ashamed of some of their music.	1.39	1.03
RP-1	7	The hair texture of members of my race is basically the same as that of whites.	1.43	1.17
RB-2	16	Dashikis and other African-type clothes are less comfortable than regular clothes.	1.53	0.84
RB-3	24	Members of my race ought to be ashamed of some of the food they eat.	1.53	1.17

school sample. This item ("Soul food is better for you than other types") had a mean value of 2.35 (s.d. = 1.45) for the former group and 4.54 (s.d. = 1.03) for the latter (t = 3.70, df = 27, $p<.001$). It is not clear, however, what this difference signifies, unless the younger group attached a more symbolic (less literal) interpretation to the item.

While the responses of the two groups were virtually identical, an examination of the direction of mean differences showed that the college sample tended to agree more with the Role Learning items than did the high school sample. This tendency probably reflects a greater awareness or education about such matters. The precise relationship between formal education and Black consciousness was not systematically investigated in the pilot study.

In general, we have interpreted the findings of this exploratory investigation into the nature of Black consciousness as consistent with the theoretical expectations; although we made some changes in the construction of a reduced instrument, the basic conceptual ideas reflected in the original scale were retained.

Black consciousness scale construction

On the basis of the results with our pilot investigation and in accord with our objective of creating a scale suitable for investigation of the relationship between Black consciousness and television identification, the scale described in previous sections was reduced to a ten-item one reflecting four basic dimensions:

A. *The extent to which racial differences are recognized rather than denied:*

Item 1: The way I look physically is basically the same as the way people of other races look. (Disagree)

Item 2: My behavior is basically the same as the behavior of people of other races. (Disagree)

B. *The extent to which racial characteristics are viewed as benefits as opposed to handicaps:*

Item 3: A dark skin is more beneficial than a light one. (Agree)
Item 4: The type of hair texture Blacks have is more beneficial than that of whites. (Agree)

C. *The extent to which racial characteristics are viewed as a source of pride as opposed to shame:*

Item 5: The most attractive nose is one that is thin and narrow. (Disagree)
Item 6: Black people ought to be ashamed of some of the food they eat. (Disagree)

D. *The extent to which differences between Blacks and whites are seen with respect to:*

Concepts:
Item 7: White people are basically alike when it comes to reacting to Blacks. (Agree)

Goals:
Item 8: The goal for Black people should be total integration. (Disagree)

Means:
Item 9: Black people should try to achieve their goals by any means necessary. (Agree)

Obstacles:
Item 10: Christianity has been a roadblock in the way of Black progress. (Agree)

Items 1, 2, 5, 6, and 8 are reversed; agreement with these items indicates low Black consciousness, while agreement with the remaining items indicates high Black consciousness.

Conspicuously missing from this reduced scale are the items about social role learning which were incorporated in our original scale. These items appeared to have more meaning for our college sample than for the high school one; since our concern was to find a generally useful scale, we decided to "err" on the side of the younger students.

Investigations employing reduced scale

The ten-item Black consciousness (BC) scale was incorporated into a general pretest questionnaire administered to subjects participating in one of the identification studies we conducted. The subjects were 50 white and 45 Black high school students recruited through newspaper advertisements. Separate administrations were conducted, with the administrator of the same race as the subjects. The purpose of including white respondents was to discover which of the items served to clearly differentiate the two groups. It was expected that all would, at least to some extent. However, it was discovered that a further reduction in the size of the scale would be necessary for experimental purposes (i.e., dividing Black subjects into "high" and "low" Black consciousness groups). Table 12 shows the means, standard deviations, and F-ratios for the data collected. Degrees of freedom are 1/93.

Table 12: Means, standard deviations and F-ratios for
ten-item black consciousness scale
(higher value = greater agreement with item)

Item	White mean	S.D.	Black mean	S.D.	F	P
1	4.50	1.37	3.11	1.75	18.72	.001
2	2.90	1.30	3.53	1.80	3.91	.05
3	2.62	1.26	2.35	1.38	0.95	ns
4	3.68	1.46	3.60	1.51	0.07	ns
5	1.56	0.97	2.06	1.45	4.06	.05
6	3.32	1.75	2.93	1.83	1.09	ns
7	3.10	1.62	4.53	1.47	20.23	.001
8	3.18	1.56	3.48	1.53	0.94	ns
9	2.50	1.70	3.95	1.69	17.39	.001
10	2.88	1.23	4.62	1.58	35.97	.001

As can be seen in Table 12, six of the ten items differentiated between the white and Black respondents. In four of these, the differences were in the expected direction (i.e., Blacks showing a higher or lower mean in the direction of "high Black consciousness"). It was expected that item 2 ("My behavior is basically the same as the behavior of people of other races") would be dissented from more often by Blacks than by whites, because that Black consciousness tends to emphasize *differences* between the races, not equality. This item may have been interpreted ambiguously by the Black students—some reading it positively, others negatively; the relatively large standard deviation would support such an observation.

One other item showed a pattern of results which was significant but in the opposite of expected directions. This item (no. 5) was "The most attractive nose is one that is thin and narrow." We expected whites to agree more than Blacks with this; however, just the opposite occurred. Our interpretation is that such beliefs about one's physical appearance are still very much evident in the Black population; while highly conscious Blacks may show less agreement than less conscious Blacks, when the two are combined (as they presumably were in our Black sample), the mean differences between them and whites either washes away or (as is the case here) actually reverses.

It is important not to impute too much theoretical significance to the results presented in Table 12. The samples are not representative of any known population, and we were more interested in developing a statistical measure which would be most sensitive to Black-white differences occurring within the general sample of 95 students along *any* race-related dimension. In most cases these dimensions revealed differences which were theoretically satisfactory in the sense that they supported the conceptions introduced earlier. In a few cases, however, these theoretical issues were thrown into question by apparent anomalies in the data. Thus, it is important to recognize that in our subsequent uses of the BC scale, when we use items which show significant differences between whites and Blacks (items 1, 2, 5, 7, 9, and 10), we are using a data-generated and not wholly theoretically-generated measure. Hopefully, additional theoretically-based research will shed more light on these issues.

In general, however, our basic theoretical position is confirmed: Blacks in our sample did perceive greater differences between themselves and whites ("other races") than did the whites. This recognition of differences *between race* and not *within race* is a crucial element. As responses to item 7 show, the Blacks think that "most whites are alike" in their reactions to Blacks.

Of course, whether or not such beliefs have an effect on the degree to which a Black student is willing or capable of identifying with television characters (and other viewers) is another question.

THE MEASUREMENT OF TELEVISION IDENTIFICATION

Freud (1922) conceptualized identification as consisting of three stages: a primary one in which a child has a close emotional relationship with his mother; a secondary one in which this relationship takes on an introjected form caused by the mother's separation; and a tertiary one in which the child begins to identify with nonparental figures on the basis of a perceived common attribute. While, strictly speaking, Freud saw this last stage as a *consequence* of identification, most of the empirical research and reconceptualization by neo- and post-Freudians views this "perceived commonality" as *constituting* identification. Much of the discussion to follow implicitly accepts this modified conception, though our statements will be made cautiously in view of Freud's original thinking.

It is also important to note, in view of later reconceptualizations, that Freud postulated identification as "consisting of" the perception of a single common attribute or quality. Some researchers—particularly the cognitive learning theorists—have made two subtle changes in Freud's conceptualization by focusing not on perception, but on *cognition*, and not on a single common attribute, but on a *number of attributes*.

The essential features of the classical Freudian conception which have implications for the measurement of identification are:

1. An element of emotional attachment or liking.
2. A perception by an observer of a model.
3. A single common characteristic shared by model and observer.

In our Experiment 1 we incorporated the three Freudian notions into our study of television identification. The particular experiment is described in detail subsequently; in the present context, it is necessary to outline the main features to illustrate how the Freudian incorporation was done.

The experiment involved two groups of male subjects, Black and white teenagers, who viewed a carefully chosen film containing three main characters—two Black and one white. The white character was a nonuniformed policeman; one of the Black characters was a uniformed policeman, and the other was portrayed as a "Black militant."

The *emotional attachment aspect* of Freudian identification was operationalized by having the subjects, after viewing the film, indicate their degree of perceived friendliness by completing a series of Semantic Differential scales (Osgood et al., 1957) containing this basic dimension. The particular polar adjectives employed included pleasant/unpleasant, friendly/unfriendly, and nice/nasty. These adjectives were selected on the basis of a list generated by students similar to those participating in the study. This operationalization was based on the assumption that a high score summated across the three adjectives would indicate high identification with that particular character.

The *perceptual aspect* of Freudian identification was operationalized merely by having the subjects view the film. While this may seem a rather obvious and/or trivial procedure, its full significance may be appreciated by contrasting it with other possible approaches: asking respondents if they identify with (or like) character X (whom they haven't seen); asking them what their favorite type of television character is; or by a number of other methods (including reading) which would exclude actual visual presentation of the model.

The *single common quality aspect* of the Freudian conception was operationalized by the selection of a film which contained characters distinctly similar or dissimilar with respect to a single characteristic—race. This single characteristic, by which viewers can perceive commonality or lack of commonality, is important because, like age and sex, it is a basic principle by which American society is organized. As such, it reduces behavioral uncertainty much more than any other single nonbiological human characteristic. Moreover, because it is a physical characteristic, race is much more likely to be crucial to a concept of identification which, in the first instance, is based on perceptual cues. The assumption here is that, *mutatis mutandis*, white viewers would identify more with white characters than with Black characters, and vice versa.

It is perhaps apparent that the three operationalizations discussed above reduced to one operation within the actual experiment. That is to say, only one aspect of identification was actually measured (the emotional attachment aspect); the other two were manipulated through a specific combination of film and subject selection. The research question, then, is: if we accept a causal model which places primary emphasis on emotional attachment as an effect (dependent variable) of the other two factors, both in isolation and in interaction, what kind of hypothesis can be generated that is consistent with Freudian notions? One general hypothesis might be: *Identification with any character, as measured by indications of emotional attachment or liking, will be highest in those instances where the viewer and the character share the common quality of race.*

Identification with the aggressor

According to the thinking of some theorists, the hypothesis given above is likely to be proven false simply because it implicitly defines identification in terms of *positive* effect. According to Anna Freud, a person can "negatively identify" with a model. Thus, instead of perceiving him as "friendly," an observer might perceive him as very "unfriendly" and yet still identify with him. How can the two conceptions be reconciled, particularly in view of the measurement procedure described above?

We attempted to reconcile this apparent contradiction by focusing attention on Anna Freud's (1947) original conceptualization which placed primary emphasis on aggressiveness. Accordingly, this dimension was also included in our Semantic Differential scales, represented by the polar adjectives aggressive/not aggressive, violent/not violent, and helpful/harmful.

Theoretically, the reconciliation occurs in the following way: identification is viewed as an "effect" having several "causes"; one cause is emotional attachment or "liking," and another is "perceived aggressiveness." Each is a *sufficient* condition for the occurrence of identification, though neither is *necessary.* Thus, both variables should be related to identification but not necessarily to each other. Thus, we can identify with a person on the basis of fear, on the basis of liking, on the basis of both, or on some other basis independent of both.

Such a conception has a disadvantage in that the "dependent" variable in this scheme—identification—ceases to have any measurement base. Thus, our basic procedure involved, as it were, a "step back" in the causal scheme to a focus on the social-structural (particularly race) variables as antecedents of identification. The "dependent" variable, identification, becomes defined as consisting of processes involving both liking and perceived aggressiveness. Identification is thus operationalized in terms of responses on relevant semantic differential scales.

Before we relate these issues directly to the research conducted, it is necessary to qualify some issues in light of Anna Freud's original conception. She was not, it seems, talking about just *any* aggressiveness. She had in mind that aggressiveness which is associated with authority figures. Translating this into the present concerns, we note that the policeman in the film stands in an authority relationship to ordinary citizens. In view of the general nature of contemporary America, the following hypothesis is thus reasonable: *Identification with policemen, as measured by indications of perceived aggressiveness, will be higher for Black viewers than for white viewers.*

It is important to note that this hypothesis refers to a particular type of identification—namely, "identification with the aggressor."

Measurement approaches derived from learning perspectives

The social learning approaches to identification are rather eclectic in nature, deriving from theories associated with a variety of perspectives. A common element of this approach—one which may also be its greatest weakness—is the equation of identification with imitation or observational learning. Our approach sought to use those aspects of the social learning theory of identification which were not dependent upon the observation of imitative behavior. (This restriction, of course, is precisely what separates social learning conceptions from those of the cognitive learning theorists; hence, the two are combined here for discussion purposes.)

One of the aspects of identification emphasized by all learning theorists is how much attention the observer pays to the model's behavior. This is important, whether one subsequently focuses on observed motoric behavior or on cognitive responses; in both cases, attention is conceived of as a necessary condition if any kind of learning is to take place.

Another reason for our focus on viewer attention is our theory about the importance of attention for the production of mass media effects. Messages which are not attended to have no effect; consequently, there is no probability of generating the necessary maintenance input to media systems.

The amount of attention paid to a model was investigated in our study by focusing on two basic domains, the *verbal* and the *nonverbal.* The former was operationalized by having viewers complete a questionnaire which contained a series of quotations recorded verbatim from the dialogue presented in the television film. These quotations were taken chiefly from the three main characters, though some fictitious ones were included in the instrument. The viewer was requested to read through the list and indicate *who said the quotation* (the sender) and *who it was* *addressed to*(the receiver). A score of one point was given for each correct answer—a possible total of two for each quotation. Since not all of the main characters had an equal number of quotations, a weighted score was devised to indicate what proportion of each character's conversation was attended to by each viewing subject. The results of this method will be discussed in more detail in our detailed description of Experiment 1.

Nonverbal attention was operationalized by having viewers respond to a series of true-false items about the nonverbal behavior of each of the main characters. A typical question was, "Character X wore eyeglasses." Again, these questions were included in an attempt to measure the extent to which observers (viewers) paid attention to specific characters.

In addition to the verbal and nonverbal measures of attention, we also included some open-ended questions in our studies. The assumption here was that the correct recall of information about the three main characters—regardless of the source of this information (the character himself, an interactant, or the television narrator)—would indicate some measurable degree of attention paid to the character.

Thus, our utilization of the contributions of social and cognitive learning theories of identification manifested itself in the incorporation of three learning-grounded measures of identification, each presumed to tap the amount of attention the viewer paid to particular characters. With respect to these three measures, the following hypothesis is relevant: *Identification with any character, as measured by any of the three attention measurements, should be highest where the race of the viewer is identical with the race of the character.* Thus, it will be noted that we find social and cognitive learning theory most relevant with respect to *measurement* concerns, while we find the *theoretical direction* offered by Freud more satisfactory.

Measurement approaches derived from social-psychological perspectives

We implicitly incorporated the major contributions of social-psychological theorists in our subject and film selection. Our subjects were chosen to represent a dominant principle of social organization—race, and our film presents models for identification which were also representative of this social structural principle.

Selection procedures for stimulus film. Most of the studies to be reported in the present paper are based on viewer exposure to a single nonexperimentally produced film. In order to understand why the particular film was chosen, it is necessary to elaborate on the relationship between social control and social communication. Mass media systems control human behavior by presenting messages which narrow the range of known behavioral patterns and which evaluate those patterns which are presented. The first stage of control is accomplished through the *selective processes* which characterize any behavior system.

When we consider these issues in relation to the social-psychological aspects of identification, it is immediately evident that there is a high correlation between one's own position in the three-dimensional social structure (race, sex, and age) and the frequency with which models who resemble one are presented in the media. Most of the identification models presented reflect the American Somatic Norm Image which Goffman (1963) has described:

> . . .in an important sense there is only one complete unblushing male in America: a young, married, white, urban, northern, heterosexual, Protestant father of college education, fully employed, of good complexion, weight, and height, and a recent record in sports. Every American male tends to look out upon the world from this perspective, this constituting one sense in which one can speak of a common value system in America. . .

The people most frequently presented in mass media are non-Black, nonfemale, and nonyouth. When those people who represent the socially unfavored position in the social structure *are* presented (recognized), they are often presented in a way which does not communicate respect —the necessary condition for legitimation. Legitimation is the fundamental energy input to social and psychological systems. Whoever controls the energy input, *ipso facto*, controls the behavior of the system in question.

When the analyst views the mass media as a social control mechanism, particularly with relationship to Black Americans, the variables he or she must attend to are recognition and respect. To be recognized is to be taken into account (Thayer, 1968), to be made relevant, to matter, to be paid attention to. To be accorded respect means to have one's behavior defined, assessed, and accounted for in the same way one would himself define, assess, and account for it (Clark, 1971c).

Black and other ethnic groups appear to go through three basic stages in terms of their presentation on commercial television. The first stage is one of virtually *nonrecognition*. American Polynesians, Puerto Ricans, Cubans, and other Spanish-speaking Americans are currently at this stage. Very few (if any) programs on the air take such people into account. They are the "irrelevant receivers."

The second stage involves some recognition, but little respect. This stage has been termed *Ridicule*. At one time (though no longer) virtually all the portrayal of Blacks on television was of the ridicule variety: *Stephen Fetchit, Amos and Andy, Tarzan*, etc. Many of these programs are still broadcast over UHF channels, but apparently they are not as prevalent as they were in the past. Instead, Asian-Americans and Chicanos (Mexican-Americans) are now presented in this manner. The Frito Bandito commerical is an example of the way television riducules Chicanos.

Asian-Americans are still typically portrayed as laundrymen, exotic sex objects, or Judo experts.

The third stage characterizes contemporary presentation of Black Americans. While the servant-master relationship is still very much present in Black portrayals (as it has always been since initial recognition), it has acquired a more structured form. In essence, the relationship has shifted to a kind of Protector/Protected relationship in which Blacks are the protectors and white society is portected. The typical Black dramatic character is thus a policeman, a detective, a spy, or some other representative of "law and order." We need to be somewhat cautious here, however; DeFleur and DeFleur (1967) have shown that virtually 70 percent of *all* television occupations are of the "law and order" variety. However, the authors did not report what percentage of these were Black in their (now somewhat dated) studies. In any case, the comparative base of interest here is not Black *vs.* white, or Black *vs.* total, but Black *vs.* Black—the proportion of Black law and order roles out of all Black roles occurring on television. Because of this predominance of policeman-type roles, we have termed this stage ***Regulation***.

This extended discussion of the types of social control exercised by television was included here simply to provide background for the selection of the particular film used in the experiments reported below. The film selection was a 1968 ***Dragnet*** film. It was broadcast commercially on network television during October 1968. Since then, it has been shown several times on UHF. Some of our studies utilized a videotape recording of the program; others used a 16mm copy distributed by MCA-TV, Hollywood, California.

The program is one-half-hour long, with intervals for three commercials (deleted in the 16mm film). The three main characters were a white middle-aged policeman ("Joe Friday"), a somewhat younger Black uniformed policeman ("Dave Evans"), and a late-adolescent Black militant ("Alex Harper"). In addition, there were several minor characters: two other nonuniformed policemen, one white and one Black; a group of white and Black adolescent men; Mrs. Dave Evans; a young Black woman, a friend of Alex's; and two shopkeepers, one white and one Black. The plot of the drama can be summarized briefly:

> On behalf of the Los Angeles Police Department, Joe Friday and his partner Bill Gannon attempt to recruit "ethnic minorities" to the police force. They call upon a Black officer (Dave Evans) to help them convince the ethnic youth (who were members of a Los Angeles graduate student union) that being a police officer is a good thing. After some initial reluctance (based on a belief that he was not a good speaker), Dave agrees to address the union, most of whose members are Black. While discussing the merits of being a police officer, Dave is interrupted and heckled several times by Alex Harper, who calls him an "Uncle Tom." Dave ignores Alex and continues to claim that "there is no such thing as a white man's law" and that the reason he joined the force was because "I wanted to do something for my country."
>
> Following this speech to the union, Dave's house is vandalized by, presumably, a group of Black youths who considered him a "traitor." Disillusioned with

> being "damned if he does and damned if he doesn't." Dave plans to quit the force, despite pleading by the two white policemen. Just before he finally does quit, however, Dave has the kind of experience he said he needed if he were to remain on the force. On a routine call, he manages to quell a "near-riot" started by the Black militant Alex. The two shopkeepers in whose store the action started said that they wanted to get a public commendation for Dave. They said that they were going to "write his captain" and praise him highly.
>
> Because Dave displayed great competence in quelling a riot and in having Alex walk away unarrested and deeply impressed by the actions of Dave, Dave is himself finally convinced that a Black man can be a credit to both races by being a policeman.

This particular film was selected because it was considered representative of the Regulation stage. Several common propaganda techniques were used, including the manipulation of *Testimony*. Two well-known football stars, O. J. Simpson and Rafer Johnson, played minor but prominent roles on the side of "law and order." Appearance of causality was manipulated by the ordering of events, contrasting the good (white suburbs) with the bad (inner city) and by the change of language (Alex's ghetto talk is transformed into proper English, like Dave's). The Somatic Norm Image was manipulated by presenting Dave looking very much like a dark-skinned white person and Alex as a very dark, short, and not-too-handsome Black. In short, the film offered a splendidly, ecologically valid rendition of important social-psychological processes at work. It was selected primarily for this reason.

CAUSALITY ATTRIBUTION: A SOCIAL-PSYCHOLOGICAL APPROACH TO TELEVISION IDENTIFICATION

The crucial element in our conception of identification—and one shared by some other theorists—is the matching process which characterizes an observer's relationship with a model. Our conception differs, however, from that of social learning theorists in that it postulates this process as a *psychological* (psyche = mind = mental) one rather than a behavioral one. In so theorizing, we explicitly align ourselves on the side of "mind" in the mind-body issue which is fundamental to all conceptions of human behavior. Of course, such mental or image matching cannot be observed by the investigators, and thus one is forced to rely upon introspective means for the inference of its existence.

We further postulate that the most important characteristic of a self-image is some kind of "wired-in" mechanism which makes it self-maintaining. We thus suggest that the underlying principle of all human behavior is a tendency for self-maintenance. This does not, of course, mean maintaining the *same* self-image; there is always elaboration of the self-image induced by the dialectic tension between the conception of the "I" (self-definition) and the "Me" (definition given by others). It is this tension which gives a dynamic system quality to the self. Hence, the matching process is one of states and not of units or entities. Some of these states, of course, are relatively stable—for example, one's race.

One of the most important aspects of this self-image maintenance involves the attribution of causality—how one views the causes of his own and other's behavior. The following proposition is central to our argument and is subjected to empirical test in Experiment 1: *An individual tends to maintain a self-image by attributing his failures or incompetencies to his environment (externally) and his successes to his dispositions or character (internally).*

We must qualify this general proposition by noting that it is characteristic of all individuals who receive some minimal amount of legitimation from the social system in which they function. Thus, only where the individual feels socially delegitimated (lacking recognition and respect) will the general causality-attribution paradigm break down. (The process of social delegitimation has been discussed in more detail in Clark [1971a].)

At first glance, one might assume that this is merely another way of saying that the proposition applies to American whites (most of whom are legitimated); if this is so, the proposition has limited usefulness in explicating general processes of psychological identification. This would be true if we were dealing with discrete states, rather than continuous ones. That is, the tendency to maintain a self-image through the type of attribution processes described above is a variable: different people are more extreme in their maintenance operation, depending on the relationship of their psychological states to the environment in which they are operating. However, the tendency to follow this type of causal attribution is considered to be uniform among all peoples.

In any case, the objection loses its potency when the proposition is seen in association with the image-matching transformation process described previously. The more general proposition is: *To the extent that a person identifies with another, he will attribute behavioral causality to that person in the same way as he would to himself; viz., manifestations of competency will be judged as internally caused and manifestations of incompetency will be judged as externally caused.*

Hence, the process of identification is not limited by the degree to which the individual conforms to the self-image maintenance pattern. Whatever the precise degree of internal attribution for success is, the person will tend to project this onto a model, if he identifies with him. Whatever deviation occurs between an individual's own pattern of causality and that which he attributes to a model, the difference is likely to be in the tendency to make him (the model) fit the proposition pattern more clearly—if, that is, the observer identifies with him. The various patterns of causality attribution are shown in Figure 3.

In order to throw Figure 3 into sharper relief, the following causality-based statements are presented as reflecting the four quadrants:

1. Black people can't do anything right because they're basically incompetent.

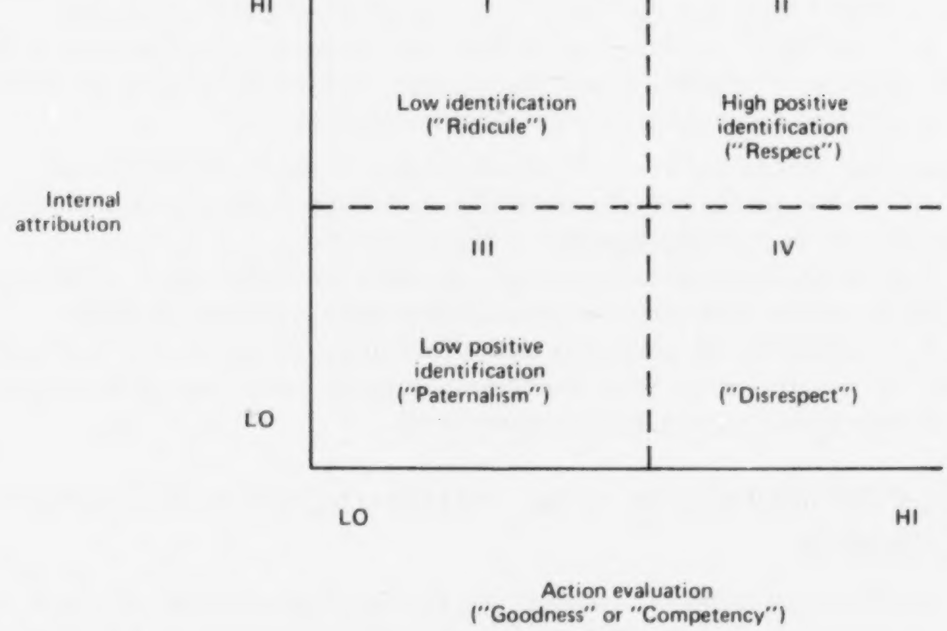

HI
LO
Internal attribution
I
Low identification ("Ridicule")
II
High positive identification ("Respect")
III
Low positive identification ("Paternalism")
IV
("Disrespect")
LO
HI
Action evaluation ("Goodness" or "Competency")

Figure 3: Hypothetical identification space

2. Black people can do anything well because they're geniuses.
3. Black people can't do anything right because they come from a poor environment.
4. Black people can do anything well because they come from a good environment.

These statements are, of course, exaggerated for effect. However, the proposition is that identification occurs to the extent that two people fall into the same quadrant both in terms of causality attribution to themselves and to the other. A "perfect" identification would thus reveal itself in a single point in the two-dimensional identification space presented in Figure 3. It thus represents an operationalization of "mutual respect" or, as many would say, "love."

Having discussed the theoretical bases of the attribution-identification measure, it remains for us to relate these to those discussed in previous paragraphs. The aspect of liking, or emotional attachment, which is central to the Freudian notion, is implicit in our more general formulation. However one defines "liking" or "emotional attachment," it could, theoretically, be defined in terms of points in our "identification space."

Seemingly absent from our conceptualization is Anna Freud's notion of "identification with the aggressor." However, it will be recalled that Freud saw this type of identification as occurring because of a lack of (or withdrawal of) emotional support. Thus, we could argue that this form of identification would be represented by the extension of our two-dimensional identification space into a three-dimensional one, adding a dimension of *attention* or *recognition* which the observer perceives he receives from the model. A high point on this third dimension would place the relationship in the "love" or "mutual respect" category; a low point would place it in the "hate" or "mutual disrespect" category.

A "love-hate" relationship, such as the concept "identification with the aggressor" implies, would thus occupy a point midway on the third (recognition) dimension, equidistant from the other two.

Thus, the conceptual measure presented does three important things:

1. It combines the contributions of Freud (with respect to emotion and attachment as central properties of identification).
2. It incorporates the variable of attention (or perceived recognition) from the various learning theoretical contributions to identification.
3. It completes the necessary and sufficient condition of total legitimation by incorporating both *recognition* (consisting of one dimension), and *respect* (consisting of two dimensions).

Operationalizing the attribution-identification measure

At this point in our discussion, we can focus more explicitly on the social-psychological foundations of our concept and indicate how it was operationalized in our studies.

Already evident, perhaps, is the close resemblance of our measure to the concept of "role taking" emphasized in Symbolic Interactionalist theory. To "take the role of" or "place oneself in another's shoes" means to extend him the same legitimation as we would to ourselves. In doing so, we are able to adopt that person's phenomenological perspective and see ourselves "in" him. In Experiment 1 we attempted to operationalize our conceptions by having viewers respond to a series of carefully chosen items relating to the characters they witnessed in the film. The questions were grouped in sets of three—a total of nine—for the three main characters in the film. Because we were particularly interested in violent behavior, nine additional items were included, to provide a total set of eighteen items.

One of the three items in each set referred to the "goodness" or "badness" of a particular action taken by a character. The viewer was required to indicate, along a seven-point scale, exactly how "good" or "successful" a particular action was.

The second item in the set asked the viewer the extent of his agreement with a statement written to provide an *internal* explanation for the character's actions. The third item (which sometimes preceded the set so as to preclude the formation of a response set) asked the respondent to indicate the extent of his agreement with a statement imputing an *external* reason for the character's behavior. Separate analyses were carried out for the internal and external measures.

In order to develop a measure which would satisfactorily handle the various theoretical aspects of the identification concept and yet be appropriate for the analysis of television identification, we decided to reduce our three-dimensional identification space to two dimensions (perceived goodness of the action and causality attribution), and assume some constant value for the amount of attention perceived. (Some indication of attention was gathered by the other measures of identification discussed earlier.)

To combine both dimensions into a single value, it was decided to take the *slope* of the measures generated by the viewer for the four responses concerning either externality or internality attributions. The diagram in Figure 4 might make this procedure somewhat clearer.

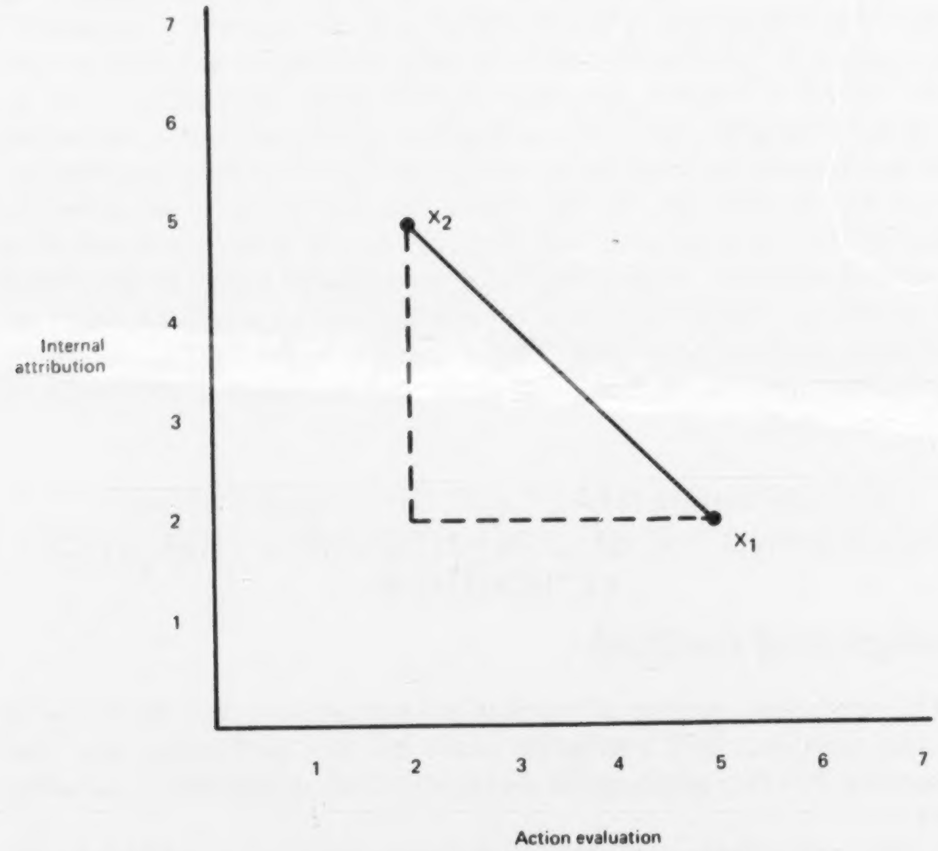


Figure 4: Calculation of slope measure

In the example presented in Figure 4, the viewer rated the first action of a particular character as +2 on the "badness-goodness" scale and as +5 on the "externality-internality" scale. His point in this two-dimensional space is presented by the X_1 occurring at the intersection of the two points. The point X_2 represents the two scores (5, 2) for the second action of the character which the viewer was asked to judge. By connecting the two points, it is possible to plot the slope of the line thus generated. Substituting the terms a_1 and b_1 for the point indicated by X_1 and a_2

and b_2 for the point indicated by X_2, the following formula gives the slope of the connected line:

$$S = \frac{b_2 - b_1}{a_2 - a_1}, \; a_2 \neq a_1$$

The value of S (our measure of identification) ranges from + 1.00 to − 1.00, with the former score representing high positive identification and the latter low positive identification (as is exemplified in Figure 4).

Another way of understanding our procedure is to consider a simple correlation matrix in which the rows ("subjects") are the various "actions" that a particular character makes and the columns ("variables") are a rating of "goodness" and "causality (internal or external) attribution." The 2 x 2 matrix generated is mathematically equivalent to the slope measure given by the above formula. Ideally, of course, more than just two actions would be examined for any character (thus resulting in a 2 x n matrix), but this would require the incorporation of additional items in the questionnaire, and, because not all actions are salient or command attention, a good deal of error variance might be generated. Of course, one could then plot a "line of best fit" by which a more sensitive slope could be calculated. This procedure would, in fact, be appropriate for shows somewhat longer than the half-hour drama we were concerned with.

AN EXPERIMENTAL TEST OF ASSUMPTIONS UNDERLYING THE SELF-MAINTENANCE THEORY OF ATTRIBUTION

Design and method

The dependent variable of interest in Experiment 1 was the extent of internal (and external) attribution made by task performers and task observers for the performer's behavior. The independent variables were:

1. The *competency* exhibited on the task (success vs. failure) by the performer;
2. The *knowledge possessed by the performer* of the observer's awareness of external factors associated with the task (i. e., factors which could be used to explain the performer's competency or lack of competency);
3. The *knowledge actually possessed by the observers* of the external factors associated with the performer's task.

All the independent variables, as well as a control condition with no observers or audience present, were crossed so as to yield a set of counterbalanced experimental conditions. To facilitate exposition, the design is shown graphically in two parts, with Figure 5 representing the uncorrelated factors.

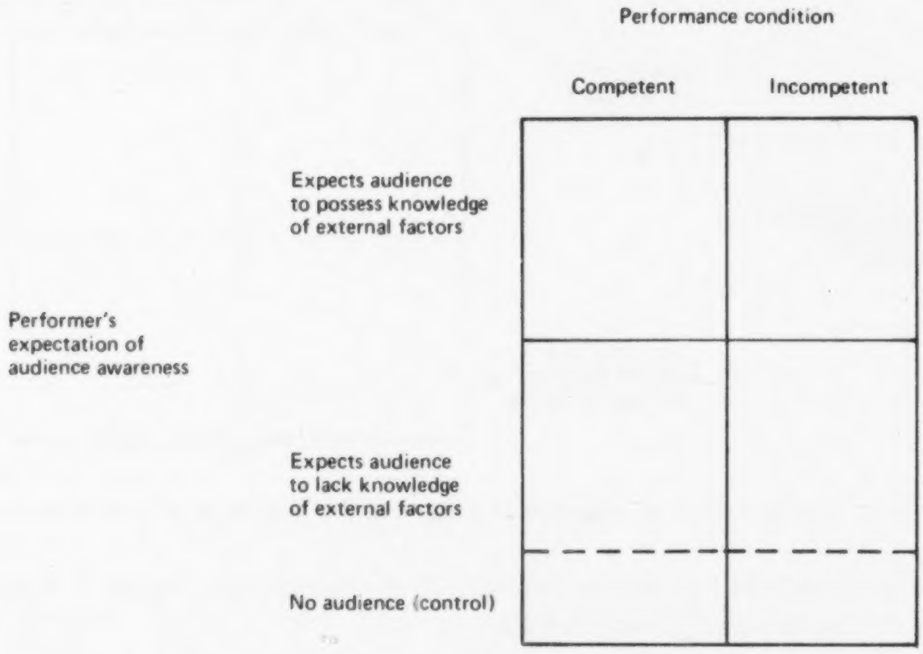

Performance condition
Competent
Incompetent
Performer's expectation of audience awareness
Expects audience to possess knowledge of external factors
Expects audience to lack knowledge of external factors
No audience (control)

Figure 5: Partial diagram of experimental design showing performance conditions

The two basic conditions represented in Figure 5 are the conditions (1) under which a performer is shown to be (a) competent or (b) incompetent and (2) under which he is led to believe that his audience (observers) (a) is aware of or knowledgeable of external constraints or (b) is ignorant of such factors.

If, as we have suggested throughout this report, individuals tend to attribute their competencies internally (to themselves) and their incompetencies externally, then there should be a main effect difference between the competency and incompetency conditions. Moreover, it should not make any difference to the performer whether he is observed or how aware such observers are of the external factors which might be claimed to affect his behavior: the levels' effect should be nonsignificant.

The design presented in Figure 5 is incomplete inasmuch as we are also interested in how the audience or observers would account for the performer's behavior. The second part of the design is presented in Figure 6.

The design in Figure 6 is, of course, directed at the behavioral responses of the observers and not of the performers. The complete design is a three-factor one (competency/performer expectation of observers' awareness/observers' actual awareness); a control condition of no

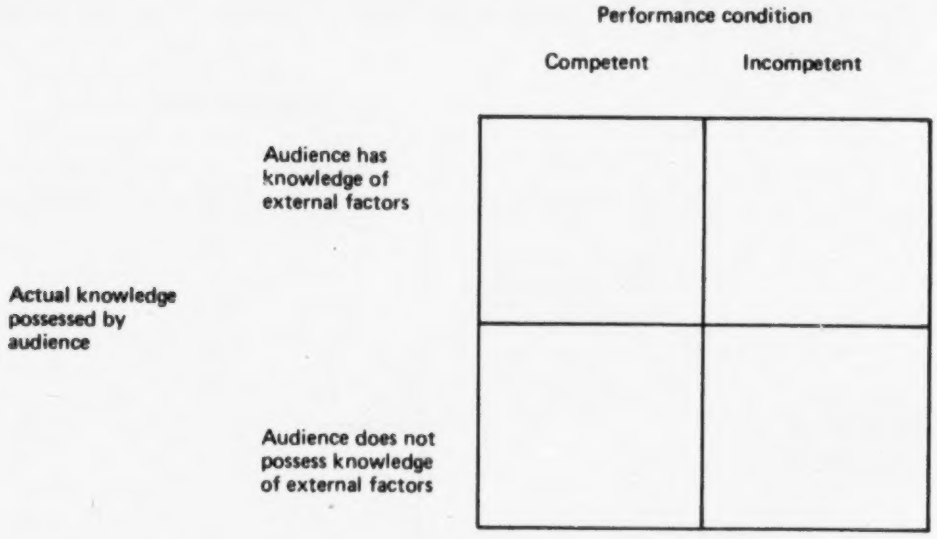

Performance condition
Competent
Incompetent
Audience has knowledge of external factors
Actual knowledge possessed by audience
Audience does not possess knowledge of external factors

Figure 6: Partial diagram of experimental design showing observer (audience) conditions

observers present is also incorporated into the design. Figure 7 shows the complete experimental design.

Method

Subjects and conditions. A total of 140 paid, white university students were run in groups of three; one was randomly assigned to the *performer* condition, one to the *knowledgeable* observer condition, and one to the *unknowledgeable* observer condition. Within the performer condition, half the subjects were assigned to the *competency* condition and half to the *incompetency* condition; within these conditions, one-third were assigned to a condition where they thought the observers were knowledgeable of external factors associated with their performance; another third were assigned to a condition where they thought the observers were not knowledgeable; and the remaining third were assigned to a no observer control condition. There were 30 subjects in each of the four main conditions and 20 in the control group.

Performance procedures. After assignment to experimental conditions, the performers were taken (one at a time) to the experimental room, where the nature of the task expected of them was described in detail. They were shown a Concept Board which contained several rows of colored lights, three response bars, and a buzzer (all rigged to a console in an adjacent room). The subjects were told that the flashing lights would signify a particular concept and that their task was to identify the concept on the basis of number, form, or color, by pressing a "yes" or "no" response bar to indicate whether the sequence of lights represented the concept in question. All subjects performed a practice trial. They were told that the trial might "facilitate" or "hinder" their subsequent performance. This was said so as to provide an external cue which could be used in attributing causality to their performance on the main task.

Competency and incompetency were manipulated by giving the performers difficult or easy tasks in the main session. In addition, they were shown interpretative charts which translated their individual scores into percentile rankings based on (fictitious) "Stanford Norms." The actual scores were manipulated by having the Concept Board preprogrammed by the experimenter.

		Performance condition	
		Competent	Incompetent
Performer expects observer knowledge of external factors	Observer has knowledge	A	A
	Observer does not have knowledge		
Performer does not expect observer knowledge of external factors	Observer has knowledge	A	A
	Observer does not have knowledge		
	No observers (performer alone)	B	B

Note: In each of the cells marked "A," the total N was 30; ten of these subjects were performers, ten were "knowledgeable" observers, and ten were "unknowledgeable" observers. There were thus 40 performers and 80 observers in the main condition. In the cells marked "B," each contained 10 (control) subjects.

Figure 7: Completed Design

Expected observer knowledge and lack of knowledge were manipulated by telling some of the performers that the two observers (who were present in the lab, and who could see the performer's face but not the pattern of lights) were aware of his practice set (i. e., that they had reason to attribute good or bad performance to external factors) and telling others that the observers were unaware of the practice set. In addition, a control group of performers completed the task with no observers present.

Observer procedures. While the performers were undergoing a practice trial in the experimental room, the observers were reading different versions of a booklet containing the task description. One version, designed for the observers assigned to the knowledge of external factors condition, described not only the main task of the performers, but also the practice set and how this might affect subsequent performance (either adversely or beneficially). The other version, which was read by those observers assigned to the *lack of knowledge* of external factors condition, described only the nature of the main task.

After reading the description of the performer's task, the observers read a description of what was expected of them as observers. Emphasized in this description was the importance of nonverbal cues and what aspects of these the observer should pay attention to while the performer was engaged in the task. The observers then joined the performers in the experimental room containing the concept board.

It should be noted again that, while the performers thought that *both* observers were either aware of his practice training or unaware of it, in fact, one of the observers was aware while the other was unaware.

Measurement of dependent variables. The major dependent variable was the extent of internal and external causality attributed for the performer's behavior. After completing the concept task, the performers and both observers completed a postexperimental questionnaire which included, among other items, the following:

1. How well do you think your performance on this test reflects your general ability at problem solving and logical thinking?
2. To what extent do you think your performance on this test may have been affected by extraneous or situational factors such as the pretraining, the presence of observers, and the experimental situation?

The former item was designed to tap the extent of *internal* causal attribution (i. e., task performance as a function of personality disposition or general competence), and the latter was designed to tap the extent of *external* causal attribution (i. e., the effect of trial practice). The items were changed slightly on the questionnaire given to observers. A 15-point scale, ranging from "not at all" to "very much," was employed. Several different statistical analyses were performed; these are discussed in more detail below.

Results

The experiment, it will be noted, was actually two experiments combined into one: one of these concerns the nature of the performer's attribution of causality, and the other concerns the nature of the observer's attribution of causality. According to our self-esteem maintenance conception of attribution, the Competency/Incompetency condition should show the greatest difference in causal attribution on the part of the performer. No main effects were predicted for presence or absence of an observer or the state of the observer's knowledge of external factors affecting performance. Also, no interactions were hypothesized to occur. In general, then, we expected performers in the Competency conditions to attribute causality more internally than performers in the Incompetency conditions. Conversely, we expected performers in the Incompetency conditions to attribute causality more externally than performers in the Competency conditions. Since internal and external factors were measured independently, the two hypotheses represent in fact two independent tests of the self-maintenance theoretical conception.

The responses of the observers in our experimental situation are not central to our basic concerns, but they are relevant to our general conception of identification. If the process of identification—at least at some level—is as pervasive as we think it is, then the following hypothesis is relevant: *Performers and actors will tend to agree more than they will disagree on their attributions of causality.*

Internality attribution responses. Table 13 shows the means and standard deviations (in particular) for the performers' internality attributions for the various experimental conditions. Figures underscored are means for the observers' responses, under conditions in which they were knowledgeable (top row) and not knowledgeable (bottom row) of the performer's practice session. A 2 x 3 analysis of variance for the performers' responses showed a significant main effect for the competency condition (F = 16.35, df = 1/54, p <.01) and no significant effects for expected observer knowledge or lack of observations. Nor was the interaction significant.

Table 13: Means of performers and observers (underscored) for responses to internality item*

	Performance condition			
	Competent		Incompetent	
No observers	9.5	(2.4)	6.6	(2.0)
Expects observers to possess knowledge of external factors	9.7	(2.3)	7.2	(2.3)
	8.8		7.6	
Expects observers to lack knowledge of external factors	9.9	(3.2)	7.2	(2.4)
	10.0		5.8	

---

*Tables 13-16 will present the means for the observers as underscored. Since the observers had no knowledge of whether the performers expected them to be knowledgeable about external factors, the means represent a collapsing over these two conditions. The meaningful figures are in the condition where the observers did or did not possess actual knowledge. The former condition is represented in the top row of Tables 13-16, and the latter in the bottom row.

To determine whether the three groups of subjects (performers, knowledgeable observers, unknowledgeable observers) differed in their internal attributions for the performer's task, a within-groups analysis of variance was done. There were no significant differences among the various respondents; only the main effect of competency *vs.* incompetency was significant (F = 25.57, df = 1/36, p <.01). None of the interactions was significant.

An additional analysis was made of just the *observer's* responses. No significant differences emerged between the observers in the two "knowledge" conditions: the only significant effect was again with respect to the competency of the performer they observed (F = 15.39, df = 1/36, p <.01). Those observers who witnessed a performance of competency were more likely to attribute causality internally than were those who witnessed incompetency. There was, however, a significant interaction (F = 7.52, df = 1/36, p <.01) between the competency and the knowledge conditions, such that the internality attribution differences were greater for those observers in the knowledgeable condition. This finding suggests that if an observer is totally ignorant of external factors affecting another's behavior, he will be even more likely to attribute demonstrated competency to the internal dispositions of the performer.

External attribution responses. An independent test of the competency-internal attribution hypothesis was conducted by having subjects respond to the second item listed earlier. It was expected that in this case those performers and observers in the *incompetency* condition would have been the greater mean value for this (external) attribution item. The mean values are given in Table 14.

A two-way analysis of variance for the performers' scores showed no significant differences among the means. The within-subjects analysis, taking into account the responses of both the performers and the observers, showed a significant effect (F = 4.84, df = 2/72, p <.05). An inspection of the means showed that the observers attributed more externality than did the performers (see comparison between the underscored and nonunderscored means in each cell of Table 14).

When the responses for the observers were analyzed separately, a significant difference occurred between the competency conditions.

Table 14: Means of performers and observers (underscored) for external attribution item

	Performance condition			
	Competent		Incompetent	
No observers	7.2	(3.7)	7.2	(2.6)
Expects observers to possess knowledge of external factors	7.3 7.6	(3.1)	5.6 9.7	(3.1)
Expects observers to lack knowledge of external factors	7.0 6.6	(3.6)	6.3 8.7	(2.8)

Consistent with an "identification" hypothesis, those observers who witnessed incompetent performances tended to attribute more external causality to the performer (F = 9.49, df = 1/36, p <.01). The observer's awareness or ignorance of such external factors had no affect on the responses.

The performers themselves did not tend to vary their attribution responses as a function of their competent/incompetent conditions. In a sense, then, the observers gave "more favorable" (in the sense that external attribution for incompetency is "better") responses to the performers than the performers gave themselves.

Further tests of performer/observer identification. In addition to the internality and externality items discussed above, subjects also answered the following two questionnaire items:

1. How well do you think the observers (performers) believe your (their) performance reflected your (their) general ability at problem solving and logical thinking?
2. To what extent do you think the observers (performers) believe your (their) performance was affected by extraneous or situational factors such as the pretraining?

In responding to these internal and external questions, subjects were thus asked to "role-play" the position (as performer or observer) of the other subjects. In a sense, then, these questions represent a more sensitive measure of identification.

Internality role-taking responses. The means for responses to the first (internal) role-taking item are presented in Table 15.

A two-way analysis of variance of the performers' responses to the internal question—in which they indicated what they thought the observers thought of their performance—showed no differences in any of the conditions.

The three-way repeated measures analysis, comparing the groups of subjects, showed a significant main effect for the between-groups competency condition (F = 17.97, df = 1/36, p <.01), with the greater internality attributed to competent performers. There was also a significant

Table 15: Means of performers and observers (underscored) for responses to role-taking internality item

	Performance condition			
	Competent		Incompetent	
Expects observers to possess knowledge of external factors	9.0	(2.7)	8.2	(1.5)
	8.4		6.5	
Expects observers to lack knowledge of external factors	9.8	(1.2)	9.0	(2.2)
	8.5		5.2	

within-groups difference, showing that the performers had a higher mean internality rating than did the observers (F = 7.46, df = 2/72, p <.01). This difference can be seen by comparing the two figures within each cell.

That the differences between the competency and incompetency conditions are due mainly to the observers' expectations about performers and not vice versa can also be seen in Table 15. This is revealed more clearly when the responses for the observers are analyzed separately. There is only one main effect, and again it is for the competency condition (F = 24.44, df = 1/36, p <.01); there were no significant interactions between observer knowledge and competency of the performer.

We find that observers, when asked what kind of attribution they think performers would make of their own behavior, tend to impute more internal causality under conditions of competency than incompetency. Interesting enough, the *degree* of internal attribution by the observers is greater than that of the performers themselves. This suggests that other people tend to think we judge ourselves "more favorably" than we actually do.

Externality role-taking responses. Table 16 presents the means for responses to the second role-taking item. None of the analysis, with the exception of a meaningless interaction in the combined knowledge conditions (see footnote to Table 13), showed any significant effects with respect to this item.

Table 16: Means for performers and observers (underscored) for externality role-taking item

	Performance condition			
	Competent		Incompetent	
Expects observers to possess knowledge of external factors	8.8 8.4	(2.3)	7.4 7.8	(2.3)
Expects observers to lack knowledge of external factors	7.0 6.9	(0.9)	6.9 8.5	(2.0)

Discussion and implications

The purpose of this experiment was to determine if our assumptions regarding the attribution process were valid. This was deemed necessary because we have placed considerable emphasis on identification and legitimation as essentially attribution-based processes. In particular we argued that the *natural* tendency is for people to attribute competency internally and incompetency externally. In contrast to other theoretical orientations, we thus grounded the process of causality attribution firmly in a self-esteem maintenance framework.

The results clearly support this conception, though responses to the externality questions were somewhat equivocal. Performers themselves attributed causality internally under conditions of competency, and people who observed them did likewise. People thus appear to tend not only to maintain their own self-esteem through such attributions, but also to "give others the benefit of a doubt" and make the same kind of attributions for them. It may well be, of course, that the homogeneity of our subject population led to a greater tendency to identify than one would normally find in the general population. It would be particularly interesting in this regard to replicate the experiment using white performers and Black observers, and vice versa.

Also of interest in the experiment was the finding that it made little difference to the performers whether they were being observed or not, or whether, if observed, they thought the observers knew as much about external factors affecting their performance as they themselves did. This suggests that people will strive to maintain a positive self-concept regardless of what other people think or regardless of whether other people are present.

This apparent disregard for observer presence also revealed itself in the performers' responses to the "role-playing" items. In accord with an internal attribution-competency hypothesis, the performers expected the observers to make the same kind of attributions as the performers themselves did (though the differences in the competency conditions were not statistically significant). What is important here is the fact that it made little difference to the performers whether or not they thought the observers possessed knowledge of factors which might affect their causality attributions; this was true for both competent and incompetent performers (see Table 15).

From the perspective of the observers, the type of attribution was almost invariably made as a function of whether the performer was competent or incompetent. In the former case, the attribution was more internal; the observers made this pattern of response both in cases where they were giving their own impressions and in cases where they were giving the responses they thought the performers themselves would give. The pattern did not vary as a function of how much information the observer actually possessed.

In conclusion, then, we can say what is important in determining the type of attribution an actor or observer will make of a given manifestation of behavior occurs primarily as a function of how successful or competent the behavior is. The more successful the behavior, the greater the probability of its being ascribed to the internal dispositions of the actor. Even under conditions where actors and observers have good reason to hold external forces responsible for the competency which is manifested, there is still a tendency for people to associate this with the actor himself. Hence, all other things being equal, people will tend to maintain a positive image for themselves and also give others the benefit of the doubt with respect to their self-images.

RACE, TELEVISION IDENTIFICATION, AND VIEWING CONTEXT

The emphasis we placed on demonstrated competency as central to an attribution theory conception of identification seems fairly well founded in light of Experiment 1 results. Having tested some of our basic assumptions in this regard, it is now possible to present some of the research findings relating our conception of identification to the process of television viewing.

Method

A total of 71 teenagers (38 white and 33 Black) participated in two sessions of the general experiment. The subjects were recruited through advertisements placed in two local newspapers—one aimed primarily at white readers, the other at Blacks. All subjects were paid for their participation.

During the first session (held separately for Black and white subjects) participants completed a background questionnaire containing several measures. Table 17 presents data describing the major characteristics of the subject populations. It should be noted that the samples do not represent any specified population and that the background data was gathered primarily for description and statistical control. As can be seen from Table 17, the differences between the two samples parallel those found in previous studies with regard to relative television viewing frequency; the Black students spent considerably more time watching television (and movies) than did the white subjects.

Table 18 presents descriptive data regarding the frequency with which specific programs were watched by the two samples. These programs were listed in the general background questionnaire and respondents were requested to indicate on a five-point scale the frequency with which they watched each program.

Of particular interest was the frequency with which "law and order" type shows were watched. These shows are underlined in Table 18. The results show that these programs were watched more frequently by Blacks than by whites in every case, with about half of them revealing statistically significant differences. These data should be interpreted with caution; as we have seen in Table 17, the general frequency of Black viewing is higher overall. Significantly, there were no differences between the two samples with respect to viewing *Dragnet*, the program used in most of our identification studies.

Table 17: Description of white and black respondents—age, class, television and movie viewing

WHITES

Variable	$\overline{X}$	S.D.		Levels of variable				
				15	16	17	18	19
Age	17.1	1.1		3%	29	34	18	16
			N=	1	11	13	7	6
				Fr.	So.	Jr.	Sr.	Other
Class	Jr.	1.0		3%	24	29	34	11
			N=	1	9	11	13	4
				Less than one hour	1	2	3	4+
Normal television viewing	2.4	1.5		11%	16	24	32	17
			N=	4	6	9	12	7
				Less than one hour	1	2	3	4+
Yesterday's television viewing	1.7	1.5		26%	18	37	3	16
			N=	10	7	14	1	6
				Less than one/month	1	2	4	5+
Movie viewing	.7	.8		50%	26	24	0	0
			N=	19	10	9	0	0

BLACKS

Variable	$\overline{X}$	S.D.		Levels of variable					
				14	15	16	17	18	19
Age	16.2	1.7		15%	27	15	9	18	12
			N=	5	9	5	3	6	4
				Fr.	So.	Jr.	Sr.	Other	
Class	Soph.	1.3		21%	33	12	24	6	
			N=	7	11	4	8	2	
				Less than one hour	1	2	3	4+	
Normal television viewing	3.4	1.4		0%	12	18	15	55	
			N=	0	4	6	5	18	
				Less than one hour	1	2	3	4+	
Yesterday's television viewing	2.6	1.9		18%	18	12	9	42	
			N=	6	6	4	3	14	
				Less than one/month	1	2	4	5+	
Movie viewing	1.2	1.2		39%	21	24	9	6	
			N=	13	7	8	3	2	

Table 18: Percent of respondents who "watch every week" selected television programs. Viewing frequency means (which appear in parentheses) are the basis for the *F* test.

	WHITE		BLACK		F
Julia	2.6%	(1.62)	24.2%	(2.82)	26.89***
Mannix	5.3	(2.43)	21.2	(2.56)	.21
Gunsmoke	0.0	(1.53)	9.1	(1.73)	1.11
The Avengers	7.9	(2.34)	6.1	(1.93)	2.74
Mayberry RFD	5.3	(1.60)	6.1	(1.91)	1.75
Bracken's World	5.3	(1.85)	3.0	(1.36)	4.94**
Bill Cosby Show	2.6	(2.74)	30.3	(3.58)	10.67***
Then Came Bronson	7.9	(2.87)	0.0	(2.04)	10.29***
Beverly Hillbillies	5.3	(1.92)	21.2	(2.78)	10.94***
Hawaii 5-0	5.3	(2.13)	30.3	(3.00)	10.73***
Here Come the Brides	2.6	(1.70)	12.1	(1.73)	.02
The F.B.I.	2.6	(1.87)	21.2	(2.62)	8.46***
Room 222	13.2	(2.70)	36.4	(3.67)	12.10***
Dragnet	13.2	(2.53)	21.2	(2.62)	.11
Adam-12	10.5	(2.42)	21.2	(2.64)	.63
Get Smart	5.3	(2.28)	9.1	(2.13)	.36
Mission Impossible	28.9	(3.30)	30.3	(3.42)	.19
The Mod Squad	2.6	(2.19)	45.5	(3.73)	32.24***
Ironside	5.3	(2.09)	15.2	(2.60)	3.90*
It Takes a Thief	13.2	(2.57)	30.3	(3.38)	7.42***

*$p < .10$, **$p < .05$, ***$p < .01$, ****$p < .005$.

Approximately four weeks after administration of the background questionnaire, all subjects participated in a viewing session under one of the following experimental conditions: (1) racially mixed group in white locality; (2) racially unmixed group in white locality; (3) racially mixed group in Black locality; and (4) racially unmixed group in Black locality. The "Black locality" was a viewing room in a public library in the predominantly Black neighborhood of East Palo Alto. The "white locality" was a similar room in the student union of Stanford University. The reason for the two physical locations was derived from a general proposition that viewing in one's "own" community would result in different identification patterns than viewing outside of one's own community. However, preliminary analysis of our results showed that this factor of locality had very little effect on the subjects. Accordingly, the cells of the design were collapsed, yielding a basic 2 x 2 (race of subject x racial composition of session) design. The average size of the experimental sessions was eight at any one time. In the racially mixed sessions, an attempt was made to equalize the number of white and Black subjects, though this was not possible in every case. However, no mixed condition ever contained a number which resulted in an unbalance of more than two.

It was unavoidable, given our random assignment procedures, that some of the students in any one session were acquainted. This was more true of the Black subjects than for the white subjects, since they were drawn from a more restricted geographical area. This did not appear to be systematically related to the experimental conditions but does provide an alternative interpretation for the differences between groups reported in subsequent chapters.

Dependent measures

The dependent measures employed were those derived from Freudian perspectives, those derived from learning theory, and those derived from social psychological theory.

Table 19: List of items included in social-psychological measure*

Item	Character	Type	Action
Joe Friday was doing a good thing in recruiting minority members for the police force	WP	E	NV
The reason Joe was involved. . .is because Joe is the type of person who would do that sort of thing.	WP	I	NV
The reason Joe was involved. . .is because his job required that he do it.	WP	Ex	NV
It was a good thing that Joe did not interfere with Dave's handling of Alex's fight with the white man.	WP	E	V
The reason Joe did not interfere was because Joe is not the type of person who interferes in other people's affairs.	WP	I	V
The reason Joe did not interfere was because of the type of situation he found himself in.	WP	Ex	V
It was a good thing that Dave quit the police force when he did.	BP	E	NV
The reason Dave started to quit the police force was because of the type of situation he found himself in.	BP	Ex	NV
The reason Dave started to quit the police force was because he is the type of person who quits.	BP	I	NV
It was a good thing for Dave to stop the fight.	BP	E	V
The reason Dave stopped the fight was because his job required that he do it.	BP	Ex	V
The reason Dave stopped the fight was because he is the type of person who likes to stop fights.	BP	I	V
It was a good thing that Alex heckled Dave at the meeting.	BM	E	NV
The reason Alex heckled Dave was because Alex is basically the heckling type.	BM	I	NV
The reason Alex heckled Dave was because of the circumstances Alex found himself in.	BM	Ex	NV
The fight between Alex and the white boy was a good thing.	BM	E	V
The reason Alex was fighting was because Alex is the type who likes to fight.	BM	I	V
The reason Alex was fighting was because of the situation he found himself in.	BM	Ex	V

*WP = white policeman; BP = Black policeman; BM = Black militant.
E = evaluation of action; I = internal attribution item;
Ex = external attribution item; V = violent action; NV = nonviolent action.

Subjects rated each item along a seven-point "agree-disagree" scale.

The social psychological measures were basically of the attribution type. On the postviewing questionnaire, subjects were requested to rate each of two actions by each of the three main characters with regard to (a) the evaluation of the action, (b) the degree to which internal forces were responsible for the action; and (c) the degree to which external forces were responsible for the action. One of the actions by each character was judged, *a priori*, as violent and the other as nonviolent. These labels should not be interpreted too literally, however; in one case, we described the policeman's breaking up of a fight as a "violent" action. The adjective describes both the action and the *context* of the action.

The two actions for each character were selected on the basis of a pilot viewing of a similar group of subjects. The subjects were asked, after viewing the film, to recall the scenes which they remembered best. Those scenes with the greatest recall frequency were judged to be the most salient and were incorporated into the questionnaire. Theoretically, the internal and external statements referring to each character should have correlated negatively. We found that this was more true of the nonviolent than of the violent actions, and, in general, we found the internal item a more sensitive and reliable index than the external item.

It will be recalled that there are two independent ways of assessing identification, using each of the attribution items. High identification is indicated, on one hand, when a subject evaluates an action highly and attributes it internally and, on the other hand, when he evaluates an ac-

Table 20: Sample of items comprising learning measure of identification

---

Verbal communication items (indicate speaker and receiver)

- "Wise investments will pay large dividends."
- "I wanted to do something for my country."
- "Everyone you help straighten out now you may not have to handcuff later on."

Nonverbal items (true-false)

- Dave Evans's wife wore eyeglasses.
- The fight took place in a butcher shop.
- Alex Harper wore a turtleneck sweater.

Open-ended items (write correct answer)

- How long had Dave been a policeman?
- What was the name of the organization Dave spoke to?
- What was the name of Joe Friday's superior officer?

tion negatively and attributes it externally. The list of statements referring to the social psychological measure appears in Table 19.

The dependent measure taken from learning theory perspectives involved responses to (a) true-false items referring to nonverbal information presented in the film, (b) open-ended items referring to general information presented, and (c) items referring to recall of specific communication exchanges. These measures were grouped with respect to each of the three main characters they involved and weighted in terms of the actual number of items and/or communication exchanges involved. Examples of such measures appear in Table 20.

The dependent measure taken from the Freudian perspective involved responses along a series of Semantic Differential scales, reflecting the dimensions of (a) friendliness, (b) violence, and (c) competency.

Identification with the white policeman

The first row of Table 21 shows the mean evaluation or assessment scores for whites and Blacks for the two actions taken by Joe Friday. It can be seen that all viewers tended to rate the two actions equally "good" along the action-evaluation scales. There are, however, significant differences in the way the two groups attributed causality to those actions. The first action was nonviolent; the Black viewers tended to account for it in terms of Friday's personality dispositions, while the whites tended to account for it in terms of social or role factors. The

Table 21: Mean assessments and attributions*

	Action 1 (nonviolent)		Action 2 (violent)	
	Blacks	Whites	Blacks	Whites
White policeman				
Evaluation	5.34a	4.94a	5.94a	6.00a
Internal	4.69a	3.69b	2.53a	2.25a
External	3.94a	4.69b	4.44a	5.19b
Black policeman				
Evaluation	3.00a	3.09a	6.00a	6.13a
Internal	2.41a	2.19a	3.97a	3.81a
External	5.75a	5.38a	5.31a	5.94b
Black militant				
Evaluation	4.09a	4.44a	3.56a	2.66b
Internal	4.56a	4.81a	4.59a	4.31a
External	3.75a	4.09a	4.53a	4.53a

*Subscripts with different letters indicate significant differences between means reading across rows for each action considered separately.

second action was "violent"; the action is again evaluated identically by the two groups, but the whites attribute causality far more to Friday's role than to his personality.

Inasmuch as both actions are evaluated relatively positively, the results indicate that the Black viewers identify more than the white viewers because they are more likely to seek internal, personal causation for the actions. The internal slope-attribution measure revealed a slightly significant difference between the two groups ($F = 3.15$, $df = 1/63$, $p < .10$), with the Blacks having the higher score. No significant differences were found with respect to the external slope measure. The difference between the two slope measures can be seen by examining each set of differences between the internal ratings for action 1 and action 2 for the whites and for the Blacks and by doing the same for the external ratings. For the Black internal ratings, the difference for the two actions is 2.16, while for the whites it is 1.44; for the Black external ratings, the difference is −0.50 and for the whites it is also −0.50.

The top of Table 22 shows the corresponding data for the mixed and unmixed viewing sessions. As can be seen, there is only one significant difference between the means, occurring with respect to the evaluation for the second action. Here the viewers in the unmixed sessions rated the action significantly higher (or better) than did those in the mixed sessions. However, because of the lack of any attribution of causality, we cannot say that one identified more with the character than the other. Both groups tended to hold forces other than the white policeman himself responsible for his particular action.

Table 22: Mean assessments and attributions*

	Action 1 (nonviolent)		Action 2 (violent)	
	Racially mixed	Unmixed	Racially mixed	Unmixed
White policeman				
Evaluation	5.13_a	5.16_a	5.69_a	6.25_b
Internal	4.03_a	4.34_a	2.44_a	2.34_a
External	4.19_a	4.44_a	4.81_a	4.81_a
Black policeman				
Evaluation	3.38_a	2.72_a	5.84_a	6.28_b
Internal	1.97_a	2.63_b	3.91_a	3.88_a
External	5.63_a	5.63_a	5.63_a	5.59_a
Black militant				
Evaluation	4.41_a	4.13_a	3.19_a	3.03_a
Internal	4.91_a	4.47_a	4.34_a	4.56_a
External	4.00_a	3.84_a	4.72_a	4.34_a

*Subscripts with different letters indicate significant differences between means, reading across rows for each action considered separately.

Significantly, no interaction (race of viewer x racial composition of session) effects occurred with respect to identification with the white policeman.

Identification with the black policeman

The middle portion of Table 21 shows the means of the various measures of judgments regarding Dave Evans, the Black policeman. As can be seen, there are virtually no differences between the white and Black viewers in evaluation and attribution of the two actions of this character. All groups tended to view Dave's actions as more externally than internally caused, and they also considered his effort to stop the fight as "better" than his efforts to recruit Blacks to the police force. The only significant difference which emerges is the tendency for the white subjects to attribute responsibility for Dave's second action (his decision to "quit the force") more to external forces. ($F = 4.43$, $df = 1/63$, $p < .05$). The data thus suggest no differences in identification with the Black policeman model between white and Black subjects.

Table 22 presents the corresponding data for the mixed and unmixed viewing sessions. The two groups evaluate the first action similarly, but they differ in the magnitude of their internal attributions. Those in the racially mixed sessions saw Dave's efforts to recruit more Blacks to the force as less internally motivated than did those in the unmixed session ($F = 3.07$, $df = 1/63$, $p < .10$). Considering the first action alone, then, there is a slight tendency for those in the racially mixed sessions to identify more with the character. With respect to the second action, we find a slightly significant difference in the evaluation of Dave's efforts to stop the fight involving Alex; those in the unmixed sessions rated it more favorably ($F = 3.29$, $df = 1/63$, $p < .10$). There were no differences, however, with regard to the attribution tendencies. The internal slope measure, combining responses to both actions, shows a significant difference between the two groups ($F = 7.37$, $df = 1/63$, $p < .01$), with those in the racially mixed sessions having the greater mean slope score.

The data thus suggest that identification with the Black policeman is higher for those who view in racially mixed sessions. Again, there were no significant interactions between race of viewer and racial composition of the viewing session.

Identification with the black militant

The bottom of Table 1 shows the means for responses to actions of Alex Harper, the Black "militant." The only significant difference occurs with respect to Alex's fight with the white teenager; the Black viewers judged it significantly more positively than did the whites ($F = 4.90$, $df = 1/63$, $p < .05$). However, the attributions for this action (as well as for the first) were identical for the two groups. Examination of the slope means showed less identification with the Black militant than for either of the other two characters, but the white and Black viewers were quite similar in this regard. (It should be noted again that, strictly speaking, we cannot compare responses among the identification models since the actions taken were different in each case.)

Table 22 shows the corresponding means for the mixed and unmixed viewing sessions. There are no differences between the two groups on any of the measures for either of the two actions.

The data did reveal one significant interaction—in response to the internal attribution item describing Alex's first action (the heckling of Dave). Those Black viewers who watched the program in an unmixed session gave a lower mean rating than those who watched in a mixed session. This difference did not emerge for the white viewers. A possible interpretation of this finding is that the Black viewers tended to identify with their white partners in the mixed session and thus tended to "stereotype" Alex by explaining his actions in terms of personality dispositions.

The drawing of definitive conclusions from the identification measures is made difficult by the failure of the data to yield consistent differences among the four viewing groups. Only with respect to the white policeman are the findings not equivocal. Here we find that Blacks identified more than whites.

With respect to the Black policeman, there were no significant differences between the two races, but there was a tendency for those in the racially mixed sessions to identify more with this character. With respect to the Black militant, no significant differences were found among the various groups.

One finding which has not been made explicit in preceding paragraphs is the general tendency of whites to attribute causality more externally and for Blacks to attribute it more internally. This is an almost perfectly consistent finding, using a variety of measures across a variety of characters and character actions. Additional analyses were performed using externality-internality tendencies as a personality correlate (the measures were taken during the pretest session); however, the variable had virtually no effect on the means presented in Tables 21 and 22. We can only conclude that there is something peculiar about the particular film (and not about the subjects themselves) which caused this distinct attribution tendency. Perhaps the Black viewers saw the entire program as accurately reflecting stereotypes of personalities, whereas the whites sought explanations outside the personalities presented.

Learning measures

Incorporated into our postviewing questionnaire were several items designed to measure the amount of attention paid by the viewer to various parts of the program. These attention measures were considered to represent the essential feature of all the theories of identification based on learning theory perspectives. These measures included: amount of correctly recalled communication interactions; the amount of recall of nonverbal material; and the amount of correct information recalled through open-ended means. The results of each of these measures is summarized in Table 23.

Communication recall (verbal measure). It is clear from Table 23 that there is a strong race effect for this variable. In every case, the white subjects had a higher mean value than the Black subjects, and no interaction effects emerged. To the extent that this measure taps some element of identification, it is clear that the white subjects identified more with all three characters than did the Black subjects. There were no significant differences between the racial viewing conditions, however.

Nonverbal (true-false) measure. Table 23 shows no significant differences between the two groups with respect to this measure. To the extent that we can compare across characters, it is interesting to note that both white and Black viewers recalled more correct information about the two policemen than they did about the Black militant. There is a slight tendency for whites to identify more with the militant than the Blacks ($F = 3.23$, $df = 1/70$, $p < .10$). There were no significant differences between the viewing composition groups, nor were there any significant interaction effects.

Open-ended measure. Significant differences emerged between the white and Black subjects with regard to the third measure of identification. The white viewers again had the higher mean scores. (For the white policeman, $F = 6.32$, $df = 1/70$, $p < .02$; for the Black policeman, $F = 7.05$, $df = 1/70$, $p < .02$). There were also significant differences between the two viewing contexts conditions, with those viewing in racially mixed sessions having the greater mean scores ($F = 4.06$, $df = 1/70$, $p < .05$). No significant interactions emerged.

Table 23: Mean values from learning measures

	Open-ended			Verbal			Nonverbal		
	WP	BP	BM	WP	BP	BM	WP	BP	BM
Blacks	.41a	.68a	*	.55a	.66a	.54a	3.4a	3.2a	2.5a
Whites	.70b	1.16b	*	.77b	.77b	.65b	3.8a	3.0a	2.8a
Mixed	.53a	1.11a	*	.69a	.76a	.62a	3.8a	3.2a	3.8a
Unmixed	.60a	.74b	*	.65a	.68a	.57a	3.5a	3.0a	2.5a

WP = white policeman (Joe Friday)
BP = Black policeman (Dave Evans)
BM = Black militant (Alex Harper)
* = no analyses performed

Means having different subscripts are significantly different from each other, reading down column for each set of comparison groups.

Taken at face value, the results derived from these learning theory measures of identification seem to indicate a much stronger pattern of identification for all characters on the part of the white subjects. This is in contrast to the findings generated by the social psychological measures, where the Black viewers tended to identify more with the white policeman than did the white viewers. With the learning theory measures, we find the whites having higher means for this character across all three indices. There is a strong possibility that these learning measures are not tapping aspects of identification so much as they are verbal fluency; this suspicion is supported by the relative lack of Black-white differences on the nonverbal measure (column 3 of Table 23).

Freudian measures

The third major contribution to the theory of psychological identification, is that developed with the context of Freudian psychoanalysis. Our operationalization of the Freudian perspectives took the form of having viewers rate along a series of Semantic Differential scales each of the three main characters. Individual scores were derived from adding each subject's rating along each of the three factors. The Semantic Differential scales were not generated *a priori* by the experimenter. An elaborate selection procedure was involved in which subjects, similar to those in our sample, generated their own list of adjectives to describe white policeman, Black policeman, etc.

Perceived friendliness of the identification model. Table 24 shows that the Black subjects rated the two policemen as more friendly than did the

Table 24: Semantic differential means for Freudian measures

	White policeman		White policeman	
	Blacks	Whites	Mixed	Unmixed
Friendly	15.97a	12.46b	13.92a	14.37a
Competency	16.59a	15.59a	15.89a	16.26a
Violent*	12.06a	11.95a	12.75a	11.23a
	Black policeman		Black policeman	
Friendly	17.41a	15.57b	16.89a	16.00a
Competency	16.68a	16.89a	17.08a	16.49a
Violent	12.76a	12.70a	13.08a	12.37a
	Black militant		Black militant	
Friendly	10.12a	8.95a	10.00a	9.00a
Competent	12.00a	11.65a	13.25a	10.34b
Violent	10.00a	8.92a	9.28a	9.60a

*A high mean for the violent dimension indicates low violence.

Means with different subscripts are significantly different from each other, reading across rows for each character.

white subjects. This is a somewhat surprising finding and is not readily interpretable. It is possible, however, that the result emerged because of the Black subjects' tendency to rate all characters, across all scales, higher than the white subjects. Equally surprising is the lack of a significant difference in ratings with respect to the Black militant, though there is a slight tendency for Blacks also to evaluate him higher than the whites.

There were no significant differences between the mixed and unmixed sessions, nor were there any significant interactions with respect to this dimension.

Perceived competency of the identification model. No differences occurred between the Black and white viewers with respect to any of the characters for this dimension. The only significant difference was between the mixed and unmixed audiences in their rating of the Black militant; those in the mixed sessions perceived him as more competent than those in the unmixed sessions.

Perceived aggressiveness of the model. Also surprising was the lack of any significant difference among the four groups in perception of violence associated with each character. As expected, the Black militant was perceived as more violent than the other characters, but this perception was, presumably, shared by Black and white subjects alike, by those who watched in mixed sessions as well as those who watched in unmixed sessions.

Our data thus show that only the perceived "friendliness" dimension reveals any significant differences among the four viewing groups. To the extent that the data are valid, they suggest that "power envy" (competency) or "identification with the aggressor" (violence) are not as sensitive measures as the "emotional attachment" (friendliness) factor which Freud himself emphasized. Of course, there is no guarantee that the Semantic Differential ratings are adequate operationalizations of the complex phenomena involved. The most justified conclusion seems to be that if one uses Semantic Differential measures of identification, the evaluative factor seems to be the best indicator. It is also worth noting that with these measures, as well as with the other two, the general finding is that Black subjects identify more than the white subjects, particularly with the white policeman.

BLACK CONSCIOUSNESS AND TELEVISION IDENTIFICATION

We expected different patterns of identification for those Black subjects who were higher in our measure of Black consciousness. In particular, we expected high consciousness viewers to identify more with the Black militant than low consciousness subjects and for the latter to identify more with the policeman than the former group.

To examine this we divided our sample of Black subjects on the basis of a median split for their Black consciousness scores (which were collected during the first viewing session). The dependent variable was the attribution-slope measure previously discussed. The mean values for the two groups for each of the three characters are presented in Table 25.

Table 25: Mean identification scores for high and low Black consciousness subjects*

	White policeman		Black policeman		Black militant	
	High BC	Low BC	High BC	Low BC	High BC	Low BC
Internal	–.37a	–.99a	–17a	–09a	.57a	.62a
External	–.19a	.01a	29a	10a	–47a	.50b

*Means with different subscripts are significantly different from each other, reading across rows for each character.

The only significant difference occurring is with respect to the external measure for the Black militant. The negative value for the High consciousness subjects shows that they tended to attribute the bad actions of this character to his environment (or situation) rather than to his personality dispositions, while the low consciousness Blacks did not tend to do this. Thus, our hypothesis suggesting higher identification for non-"law and order" characters by high Black consciousness subjects is supported.

Summary of results of identification measures

In any effort to determine the extent to which the Freudian, learning, and social psychological measures were tapping the same or distinctive components of identification, a factor analysis of all the dependent variables was carried out.

The results of this analysis showed that the various items loaded more on the basis of item format than on the basis of character judged. This was particularly true of the learning measures; the open-ended measures loaded together across all characters and were distinct from the true-false measures, which loaded together on another factor. The analysis thus confirmed our earlier suspicions that the learning measures were not indicating identification with particular characters so much as a general ability to retain and reproduce information provided in the film.

The Freudian (Semantic Differential) measures were much more sensitive indicators and resulted in unique factors for each of the three main characters. There may, however, have been a methodological artifact involved here as well, since the measures did not correlate very well with other identification measures. The competency, friendliness, and violence variables tended to correlate highly with regard to a particular character, but the cluster itself showed rather low correlations with both the learning measures and the slope-attribution measures.

In contrast, the slope-attribution measure was spread throughout the various factors; it showed no unique structure of its own. There are two possible reasons for this: (1) the variance associated with this measure may have been so minimal as to preclude a significant loading, or (2) the measure may have been tapping a unique element of identification which was common across all measures. This latter interpretation is more satisfactory in view of our basic theoretical expectations, but more research is needed before unequivocal conclusions can be drawn.

Theoretical summary. Table 26 shows a descriptive summary of the main findings of the research we have described. The various measures —particularly the learning ones—appear to be tapping different aspects of the identification phenomena, and our results were not in every case consistent. Some of our statistically significant results are likely to be the result of chance factors, particularly in the studies involving a rather long postviewing questionnaire. The tentative conclusions are presented below more as suggestions for further research than as definitive statements concerning the nature of race and television identification:

Table 26: Summary of identification findings

	White policeman	Black policeman	Black militant
Measure 1 (Attribution)	B > W	nd	nd
Measure 2 (Evaluation)	B > W	B > W	nd
Measure 3 (Learning)	W > B	W > B	W > B

B = Black subjects
W = white subjects
> = greater than in identification score
nd = no significant differences

1. Black subjects, on the whole, tend to attribute responsibility for actions engaged in by television characters to internal factors, while whites, on the whole, tend to attribute them to external factors.
2. Because most of the actions taken in this particular film were "good," the result was a greater degree of identification with these characters on the part of Black subjects.
3. Measures of identification using evaluative scales show that Blacks identify more than whites.
4. Measures of identification using learning scales show that whites identify more than Blacks.

5. There are no consistent patterns showing that Black subjects identify more with Black characters, nor whites with white characters.

6. Racially mixed audiences tended to identify more than racially unmixed audiences, but this was not with respect to any particular character.

7. Black subjects who are "conscious" of being Black tend to identify more with Black characters, particularly those who are presented in contrast to "law and order" models.

8. Our conclusions are based only on an intense analysis of a single film which was selected because of its "typicality" along theoretically relevant dimensions.

SUMMARY AND CONCLUSIONS

We will not attempt here to give an exhaustive summary of the various theoretical and empirical conclusions discussed. Instead, we will present, in statement form, some of the major points we would like to emphasize by way of conclusion.

1. All systems produce outputs. The outputs of mass media systems are messages and images.

2. The images of nonconsumers (e.g., Blacks and other American ethnic minorities) are unfavorable in comparison to the images presented of the consumers.

3. Psychologically, this enables the consumers (who are mostly white) to gain a sense of self-legitimacy or self-esteem by comparing themselves with the nonconsumers (who are mostly nonwhite).

4. Since legitimacy (self-esteem) is a necessary input for the adequate function of both social and personality systems, the mass media's refusal to grant this to nonwhite Americans helps to produce personality destruction.

5. Such psychological destruction is what can be defined as "television violence."

6. The legitimizing function of the mass media reflects itself in the presentation of models with whom receivers can identify.

7. The mass media have tended to present nonwhite models as (a) nonexistent; (b) objects of ridicule; or (c) objects of political regulation. Such models are difficult for nonwhites to identify with.

8. They have been particularly difficult for people who are high in "race consciousness."

9. Through dialectical system processes, "race consciousness" is raised by virtue of such people's exposure to mass media models with whom they can not identify.

10. The social violence engaged in by such "race conscious" people is an effect caused, in large measure, by the social violence engaged in by the mass media through its delegitimation of nonwhite Americans.

11. A large proportion of the identification models currently presented to Black Americans are those involving "law and order" occupations.

12. Studies using various measures of television identification show that in one typical "law and order" program:

   a. Blacks tend to identify more than whites with the "law and order" models.
   b. The pattern of Black identification with television models changes as a function of "race consciousness." This is more true for Black television models than for white television models.

13. Further research, both theoretical and empirical, is needed in order to further specify the dynamics of Black television identification.

REFERENCES

Berrien, F.K. *General systems and social systems.* New Brunswick, N.J.: Rutgers University Press, 1968.

Billingsley, A. *Black families in America.* Englewood Cliffs, N.J.: Prentice-Hall, 1969.

Clark, C. *The concept of legitimacy in black psychology.* In Epps, E. (Ed.) *Race relations: current perspectives.* Winthrop, 1971, in press (a).

Clark, C. Social change and the communication of legitimacy: the case of dispute settlement. *Journal of the Developing Areas,* 1971, **5** (4), 577-88, in press (b).

Clark, C. Competency and legitimacy as organizing dimensions in the Black self-concept. Paper presented to the Self-Concept work group, Social Science Research Council, Chicago, 1971 (c).

Clark, C. The televised portrayal of ethnic groups: data and theoretical considerations. Unpublished manuscript, Stanford University, 1970.

DeFleur, M, and DeFleur, L. The relative contribution of television as a learning source for children's occupational knowledge. *American Sociological Review,* 1967, **32**(5), 777-89.

Easton, D. *A framework for political analysis.* Englewood Cliffs, N.J.: Prentice-Hall, 1965 (a).

Easton, D. *A systems analysis of political life.* New York: Wiley, 1965 (b).

Festinger, L. A theory of social comparison processes. *Human Relations,* 1954, **7**, 117-40.

Freud, A. *The ego and the mechanisms of defense.* New York: International University Press, 1947.

Freud, S. *Group psychology and analysis of the ego.* London: Hogarth, 1922.

Goffman, E. *Stigma.* Englewood Cliffs, N.J.: Prentice-Hall, 1963.

Heider, F. *The psychology of interpersonal relations.* New York: Wiley, 1957.

James, W. *Psychology: the briefer course.* Edited by Gordon Allport. New York: Harper & Row, 1961.

Kelley, H. Attribution theory in social psychology. In Levine, D. (Ed.) *Nebraska symposium on motivation.* Lincoln: University of Nebraska Press, 1967.

Klapp, O. *Heroes, villains, and fools.* Englewood Cliffs, N.J.: Prentice-Hall, 1962.

Laing, R. D. *The divided self.* London: Tavistock Publications, 1960.

McCall, G., and Simmon, J. *Identities and interactions.* Glencoe, Ill.: Free Press, 1966.

Mead, G. *Mind, self and society.* Chicago: University of Chicago Press, 1934.

Newcomb, T. An approach to the study of communicative acts. *Sociological Review,* 1965, **60,** 393-404.

Osgood, C., et al. *The measurement of meaning.* University of Illinois Press, 1957.

Robinson, A., et al. (Eds.) *Black studies in the university.* New York: Bantam, 1969.

Rotter, J., et al. Internal versus external control of reinforcement: a major variable in behavior theory. In Washburne (Ed.) *Decisions, values, and groups,* Vol. 2. New York: Macmillan, 1962.

Taylor, O. *Black English.* Unpublished manuscript, Center for Applied Linguistics, Washington, D.C., 1970.

Thayer, L. On communication and change: some provocations. In Shoemaker, F.F. (Ed.) *Proceedings of the 11th Annual Institute in Technical and Organizational Communication.* Colorado State University, 1968.

Thomas, C. (Ed.) *Black perspectives in psychology.* Glencoe, Ill.: Free Press, 1971.

Social Class and Racial Differences in Children's Perceptions of Television Violence

Bradley S. Greenberg and Thomas F. Gordon

Michigan State University

Questions about the effects of observing violent or aggressive behavior have been given considerable attention; investigators have ranged from social scientists to theologians to U.S. Senators. Much literature has examined the likelihood that a child will behave aggressively after viewing an aggressive model. Almost as plentiful are the reviews of behavioral studies which attest to the quantity of work in this area (Flanders, 1968; Goranson, 1969; Weiss, 1969).

This literature is virtually void of empirical efforts to determine the meaning of the stimulus for the youngster or his attitudes toward such content. What is "violent" or "aggressive" has been specified primarily by the researcher or commentator, and use of these terms has differed widely. For example, Bandura (1963) contends that intent is an important element of aggression, whereas Buss (1961) holds that we must exclude intent and deal with behavioral acts. Where some investigators posit an aggressive drive (Kahn and Kirk, 1968), others argue for motivations (Epstein, 1962). These explications are little concerned with the viewers' definitions or perceptions of the content.

Content analysis research which attempts to identify and count the types, degrees, and extent of violence in entertainment programming also generates a single-sided definitional scheme (Gerbner, 1969; Greenberg, 1969). Weiss (1969) succinctly pointed out the lack of receiver-oriented studies of media violence in his recent review:

> . . . .there is a total lack of information concerning the subjects' definition of the experimental situations and the meanings or interpretations they gave to the movie or the behavior of the models, or concerning their reactions during the observation of the model or the movie. In the absence of such knowledge . . . any facile assumption about the viewers' reactions and interpretations should be viewed with considerable caution.

Weiss called for studies which would deal with perceptions of media violence and would compare the perceptions of different subgroups of viewers. Such studies, he contended, could yield evidence necessary for more accurate interpretation and prediction of the effects of observing media violence.

The present study deals with these issues.[1] Specifically, we examined young boys' perceptions of selected television program scenes of differing kinds and degrees of violence. A central focus of the study was the manner in which attitudes toward a given vignette differed when racial and socioeconomic factors were considered. The questions became: how much, if any, violence was judged to be present in the scene; how acceptable was the depicted behavior; how did these perceptions vary among subgroups of viewers?

Some previous laboratory experiments on perceptions of violence are pertinent to our study rationale. In those studies, the typical approach has been to use a stereoscope viewer to present a different image exposure to each eye. Through binocular fusion, these separate pictures merge. These studies have examined predictors of what is perceived when a violent image is presented to one eye and a nonviolent image to the other.

Two major influences on what the individual sees in such experiments —the age and the background of the viewer—have been isolated (Toch, 1961; Reif, 1967; Moore, 1966). For example, among children in grades 3-13, age is positively related to the increasing tendency to perceive violence. Especially important for the present study is evidence about the influence of the individual's background on his perceptions. Reif (1967), working with institutionalized male delinquents, found that those with a background of aggressive behavior and with current aggressive habits perceived relatively little violence in his stereoscope experiments. They saw less violence than did delinquents without a history of aggressiveness and/or without current aggressive tendencies. In other words, the youngsters most directly exposed to and involved with aggression were less likely to see the stimulus as violent.

The manner in which people become involved with violence appears to be critical in determining their perceptions of it. For example, professional critics of television see more violence in "violent" programs than does the general public (Greenberg and Gordon, 1971). Toch and Schulte (1961) found a similar sensitivity toward seeing violence among a group of police trainees. Those who had recently completed training were more prone to see a violent image in a stereoscope than were incoming trainees.

The present study contrasts this tendency to see violence by those who are being professionally trained and rewarded for doing so with the tendency among those for whom violence has not been rewarding but for whom it is more commonplace, like Reif's (1967) delinquent boys. For the latter, given greater frequency of direct exposure to various forms of "real-life" violence, where seeing violence is not rewarded, the expectation is that less television violence will be perceived. When violence is judged to be present, its intensity should be assessed more moderately. Gross acts of mayhem will surely always be described as containing some degree of violence, but the central issue is how much?

Ample evidence indicates that physical aggression is a common aspect of the daily lives of youngsters from lower-class environments. In low-income homes, physical punishment is used more frequently as a mode of correction than verbal approaches (Sears, 1961; Chilman, 1965; Gans, 1962; Moles, 1965). The environment outside the home is also more likely to be hostile for the low-income child, especially if he lives in a ghetto area (Clark, 1965; U.S. Government, 1968).

This evidence contributes to the rationale that greater exposure to real-life aggression manifests itself in greater tolerance for aggressive behavior, whether real or mediated. The youngster who has been thus exposed to more aggression may become inured or sated by his more frequent exposures. For these reasons, the following hypotheses were tested:

H1: The less advantaged youngster will perceive less violence in a given segment of television violence than will a middle-class youngster.
H2: The less advantaged youngster will judge mediated violence as a more acceptable mode of behavior than the middle-class viewer.
H3: The less advantaged youngster will see television violence as more real than will his middle-class counterpart.

For parallel reasons, the less advantaged youngster should be more attracted to programs high in action and excitement, of which the violent programs are the principal available examples. Indeed, given the greater amounts of time lower-class children spend with television (Greenberg and Dervin, 1970), such physical action may be necessary to arouse the viewer to attentiveness. Thus, these hypotheses were tested:

H4: The less advantaged youngster will judge violence as more enjoyable to watch than will a middle-class youngster.

H5: The less advantaged youngster will testify to more self-arousal from television violence than will a middle-class youngster.

Although these hypotheses would be the major emphasis of this study, we concurrently tested certain subhypotheses about the weaponry of violence. Himmelweit, Oppenheim, and Vince (1958) suggested that children watching violent programs are aroused to differing degrees depending on the types of weapons involved. More recently, Berkowitz (1964) demonstrated a positive correlation between symbols of aggression like guns and knives and the likelihood that aggressive behavior would be elicited in the presence of such symbols. Generalizing from these observations, we posited the following:

H6: Violent scenes with weapons will be perceived as more violent than violent scenes without weapons.

H7: Violent scenes with weapons will be perceived as less realistic than violent scenes without weapons.

METHODS

Preteen males were shown a variety of kinds of television violence mingled with nonviolent sequences. Testing was done with groups of four to six boys in a room in their public school. Attitudinal responses were obtained in terms of several sets of verbal scale items. The socioeconomic status and race of the boys varied.

Video materials

Twenty-three hours of prime time television containing 24 programs were taped to obtain material representative of the array of television violence in current programming. All programs were taped between 7:30 and 11 p.m. over a three-week period, February 2-23, 1970. (The complete set of programs is listed in Appendix A.) Programs taped were chosen on *a priori* expectations of violent content if they were new that season and from previous data if they were from a returning series (Gordon, 1969).

From this recorded material, all individual violent sequences were edited onto a master tape. This 45-minute tape contained 75 separate scenes of violence which varied in length from five to 120 seconds. All violent sequences were scenes in which characters physically harmed themselves or another person (e.g., hitting or shooting), overtly intended such harm (e.g., shooting but missing), or physically damaged some inanimate object (e.g., smashing furniture). Scenes of yelling or shouting were also recorded as examples of verbal aggression. The latter were so few in number as to preclude their further examination in this study.

In this sample of content, the three major types of physical violence evident were property destruction, physical assault against others, and accidental or intended death. Two scenes of violence were chosen from each of these three major types; two scenes of more idiosyncratic violence, a suffocation and a fire death scene, were also included. Two stimulus tapes were created from these scenes, adding practice and control scenes. Each version had one scene of each violence type, a fourth violent scene, the same practice scene, and the same two control scenes—for a total of seven scenes.

The violent scenes were randomly ordered in one tape; the other was constructed with a parallel order, so that order of presentation was constant for all subjects. Capsule descriptions of the scenes in each version are presented in Figure 1 in their order of appearance. More complete descriptions are in Appendix B.

Subjects

All subjects came from public schools in Kalamazoo and Grand Rapids, Michigan. In each city, the superintendent's office designated the primary socioeconomic class makeup of each school in the system and provided racial census data. Schools were then selected which provided adequate numbers of subjects in four economic and racial categories: lower-class white, lower-class black, middle-class white, and upper-class white.

In all schools, fifth grade boys were used. Parental nonpermission slips were distributed, and less than one dozen nonpermission slips were received. This yielded 325 fifth graders for the data analysis: 89 lower-

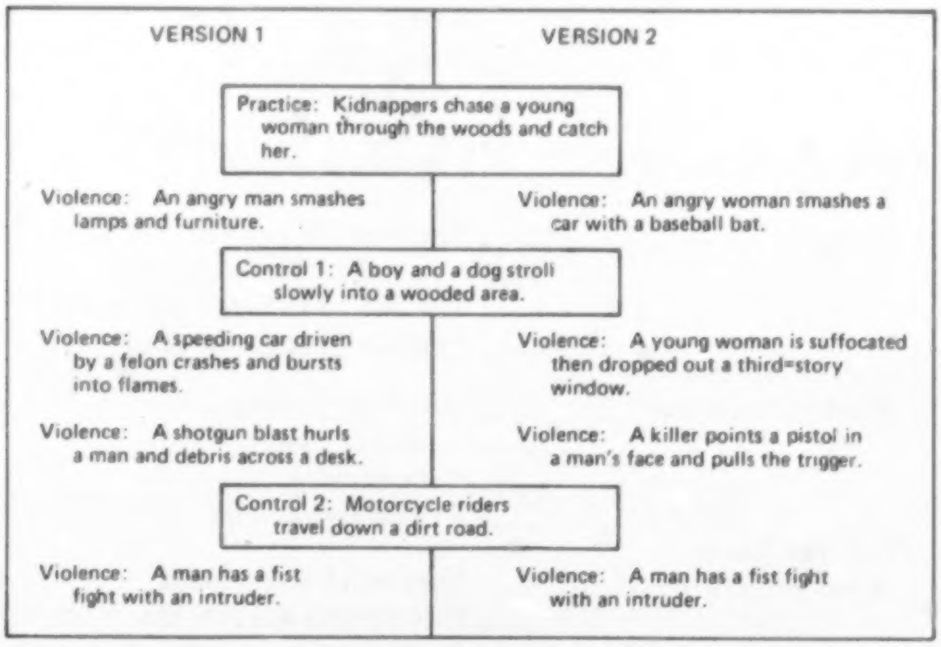

VERSION 1
VERSION 2
Practice: Kidnappers chase a young woman through the woods and catch her.
Violence: An angry man smashes lamps and furniture.
Violence: An angry woman smashes a car with a baseball bat.
Control 1: A boy and a dog stroll slowly into a wooded area.
Violence: A speeding car driven by a felon crashes and bursts into flames.
Violence: A young woman is suffocated then dropped out a third-story window.
Violence: A shotgun blast hurls a man and debris across a desk.
Violence: A killer points a pistol in a man's face and pulls the trigger.
Control 2: Motorcycle riders travel down a dirt road.
Violence: A man has a fist fight with an intruder.
Violence: A man has a fist fight with an intruder.

Figure 1: Scenes in film versions

class blacks, 89 lower-class whites, 90 middle-class whites, and 57 upper-class whites.

Instrumentation

The test instrument was designed to assess the subject's evaluation of the stimulus scenes along five specific attitudinal dimensions. These *a priori* dimensions dealt with how much, if any, violence was perceived in each scene, how acceptable the demonstrated behavior was judged to be, how much the subject enjoyed viewing each content segment, how realistic the content was perceived to be, and how much self-arousal was felt. These dimensions were developed to correspond to the hypotheses stipulated earlier.

An early form of the instrument was pretested, with the same content segments, on 60 students in a different public school system. In addition, public school teachers were asked to evaluate procedures, forms, and wording. Findings and suggestions were incorporated into the final instrument.

Three items were constructed for each of the five attitudinal dimensions, and the same 15 items were used for all seven stimulus scenes. Here are sample items for each dimension:

Degree of Violence
Was what you saw .... ____ Extremely violent
____ Very violent
____ Pretty violent
____ Not very violent

Acceptability of the Behavior
Is it .... ____ Very right for people to be this way
____ A little right
____ Not very right
____ Not right at all for people to be this way

Liking of the Content
Was what you saw a show like .... ____ You really like to see
____ You sometimes like to see
____ You don't like to see very much
____ You don't like to see at all

Degree of Arousal
What you saw was .... ____ Not very exciting
____ A little exciting
____ Very exciting
____ Extremely exciting

Perceived Reality
What you saw was ...... ____ Very much like real life
____ Pretty much like real life
____ Not much like real life
____ Not at all like real life

The complete set of items is in Table 1, together with data on the empirical verification of the dimensions.

Procedures

Testing was conducted in April and May, 1970. Each school provided a room large enough for the videotape equipment and for four to six children seated in front of a television set. Group size was limited so that each subject would be close to the 21-inch screen and interaction among them would be minimal.

The boys were told that we wanted their reactions to scenes from regular television programs, that this was not a test and would in no way affect their classroom evaluation. They made no personal identifications on the instrument. Booklets were coded for race and video treatment condition after the boys left the viewing room.

Subjects first viewed the practice scene. The experimenter completed two or three items with them to clarify how they were to proceed. The boys completed the remaining items for the practice scene and were questioned about difficulties with words or procedures. When they understood the items and procedures, each child was asked if he wished to continue. (Of 329 subjects, four declined to continue.) The subjects were then shown the six remaining scenes and rated each scene immediately after viewing it. On the average, they took 25-30 minutes to view and rate all seven scenes.

The boys were asked not to talk to their classmates about what they had done until everyone had participated. Teachers were asked not to discuss the children's experiences with the class until testing in that school was completed.

RESULTS

Four major analyses were completed: (1) a factor analysis of the test items; (2) a comparison of racial and social class differences in response to the stimuli; (3) an examination of differences among kinds of violence; and (4) a check on relative perceptions of the control and experimental scenes.

Item analysis

Three test items were designed to tap each of five attitudinal dimensions. To determine if the *a priori* allocation of items to these categories had empirical support, item responses were intercorrelated and then submitted to a principal axis factor analysis with varimax rotation. A summary of the factored results is in Table 1.

This procedure yielded three major factors and a minor one. One major factor was that of perceived *violence* in the stimuli. Four items

Table 1: Factor items

			Percentage of total variance
Factor 1. Perceived violence			17 percent
ITEMS:		Item loadings	
Were the people	____Not very angry ____A little angry ____Very angry ____Extremely angry	.64	
Was what you saw.	____Not very violent ____Pretty violent ____Very violent ____Extremely violent	.58	
Was what you saw	____Not very serious ____A little serious ____Pretty serious ____Very serious	.71	
Was what you saw	____Not very cruel ____A little cruel ____Pretty cruel ____Very cruel	.58	
Factor 2. Perceived acceptability			17 percent
ITEMS:		Item loadings	
Is it	____Very right for people to be this way ____A little right ____Not very right ____Not right at all for people to be this way	.84	

Table 1: Factor Items—Continued

Factor 2. Perceived acceptability—Continued			Percentage of total variance
ITEMS:		Item loadings	
Was what you saw	____ A very good thing to do ____ A pretty good thing ____ A pretty bad thing ____ A very bad thing to do	.82	
Is it	____ Very nice for people to act like this ____ Pretty nice ____ Not very nice ____ Not nice at all for people to act like this	.81	
Factor 3. Professed enjoyment			25 percent
ITEMS:		Item loadings	
What you saw was	____ A very good thing to watch ____ A pretty good thing ____ A pretty bad thing ____ A very bad thing to watch	.77	
What you saw was	____ A very funny thing to see ____ A pretty funny thing ____ A pretty sad thing ____ A very sad thing to see	.77	
Was it	____ A wonderful show ____ A pretty good show ____ A pretty bad show ____ A terrible show	.80	
Does what you saw	____ Make you feel like laughing a lot ____ Make you feel like laughing a little ____ Not make you feel like laughing very much ____ Not make you feel like laughing at all	.70	

Table 1: Factor Items–Continued

Factor 3. Professed enjoyment–Continued			Percentage of total variance
ITEMS:		Item loadings	
Was what you saw a show like	____ You really like to see ____ You sometimes like to see ____ You don't like to see very much ____ You don't like to see at all	.79	
Factor 4. Perceived reality			7 percent
ITEM:		Item loadings	
What you saw was	____ Very much like real life ____ Pretty much like real life ____ Not much like real life ____ Not at all like real life	.87	
The following items were too impure to assign to a single factor:			
What you saw was	____ Not very exciting ____ A little exciting ____ Very exciting ____ Extremely exciting		
Was what you saw	____ Very much for fun ____ Pretty much for fun ____ Not very much for fun ____ Not for fun at all		

loaded primarily on this factor and accounted for 17 percent of the total variance. Three of the items had been designated as violence perception items. The fourth dealt with the judged seriousness of the stimuli.

A second major factor was perceived *acceptability*. The three items originally constructed for this dimension loaded together and explained 17 percent of the total variance.

The third major factor was professed *enjoyment* of the scenes. Five items loaded on this factor and accounted for 25 percent of the total variance. These included the three items designed for this dimension, plus two items originally conceived of as arousal indices—feelings of laughter and the humor of the scenes. The latter two perceptions were measured as part of the subjects' overall enjoyment of the television content. No arousal factor emerged.

The fourth factor was but a single-item assessment of the perceived *reality* of the content. (Two items had impure loadings and were omitted from subsequent analyses.) These four factors accounted for 66 percent of the total variance. We posited five factors in the original instrument, including arousal and reality factors; three strong factors—perceived violence, acceptability, and enjoyment—emerged from the factor analysis.

Social class and racial differences

The study rationale hypothesized that attitudinal responses to the television violence would order from upper- and middle-class white males to lower-class whites, to the single group of lower-income black males tested. Basically, class difference responses were predicted, with a single racial comparison expected to intensify such differences.

For each of the four dimensions of judgment, item scores were summed. Given two versions of the experimental stimuli[2] and the repeated measures within each version, a Friedman two-way analysis of variance by ranks was used to test the perception hypotheses across the eight scenes of violence (Siegel, 1956). In addition, subanalyses for each factor: (1) compared the group ratings within each violent scene, (2) compared ratings across the groups for all violent scenes collapsed, (3) compared the combined lower-class groups with the combined middle- and upper-class groups, and (4) compared the scene ordering within each factor. Violent scenes are compared with nonviolent scenes in a subsequent section.

Perceived violence. Table 2 presents the mean values of perceived violence. Across the four groups, the analysis of variance by ranks identified one major difference in perceptions of the eight violent scenes. From the sum of ranks, the origin of this difference was the viewer's race. The black lower-class group saw significantly less violence across all scenes ($p<.001$) than all other groups.

Table 2: Perceived violence†

Scene††	Social group Black Lower	White Lower	White Middle	White Upper
Pistol killing**	12.10 (1)	14.15 (4)	13.73 (3)	13.32 (2)
Shotgun killing	12.76 (1)	13.29 (2)	13.33 (3)	13.39 (4)
Suffocation killing	12.35 (1)	13.29 (3)	13.35 (4)	13.19 (2)
Death by fiery car crash	12.35 (1)	13.04 (3)	13.43 (4)	12.92 (2)
Smashing car	11.95 (1)	13.27 (4)	13.17 (3)	12.77 (2)
Fist fight No. 1**	11.18 (1)	13.21 (3)	13.95 (4)	13.04 (2)
Fist fight No. 2**	10.85 (1)	12.42 (2)	12.73 (4)	12.48 (3)
Smashing furniture*	10.96 (1)	12.10 (2)	12.33 (3)	12.73 (4)
Sum of ranks:	(8)	(23)	(28)	(21)
		($X^2r = 16.35, p < .001$)		
Mean ratings across scenes:	11.81	13.08	13.25	12.98

†The larger the mean, the more violence; the higher the rank, the more violence.

††Mean differences for individual scenes significant by one-way analysis of variance: *$p < .05$; **$p < .001$.

The subanalyses confirmed this interpretation, First, in a one-way analysis of variance among the four class levels for each scene, four scenes produced significant differences (see asterisks in Table 2). In each case, the difference was due primarily to lesser violence perceived by the black lower-class respondents. Second, the eight violent scenes were collapsed and a one-way analysis of variance was computed across the four class-race group means (bottom row of means in Table 2). The groups differed significantly, the variation due primarily to the lower ratings of the black lower-class group ($p<.001$). Third, using these collapsed means, the combined lower classes were compared to the combined middle and upper classes by t-test. The results were significant, with the lower-class boys seeing less violence ($p<.01$).

The Kendall coefficient of concordance was computed to determine how similarly the four groups ordered the eight scenes by degree of perceived violence. The result was a .52 correlation among the four groups in ordering the scenes ($X^2 = 14.59$, $df=7$, $p<.05$). Table 2 lists the scenes in ranked order from most to least violent. Rated most violent were the pistol and shotgun killings; rated least violent were the fist fight and furniture-smashing scenes.

Perceived acceptability. Table 3 presents the data for this attitude factor. The analysis of variance by ranks was significant across the four groups for the eight violent scenes ($p < .001$). The difference was one of both economic class and race, with race accentuating the income difference. In six of the eight violent scenes, the lower-class boys were most likely to call the behaviors they observed acceptable; in seven of eight, the black youths were most extreme in acceptance.

Table 3: Perceived acceptability†

	Social group			
Scene††	Black Lower	White Lower	White Middle	White Upper
Fist fight No. 2***	7.85 (1)	9.44 (2)	9.85 (3)	10.19 (4)
Smashing furniture	9.29 (1)	9.79 (2)	10.50 (4)	10.00 (3)
Fist fight No. 1***	8.94 (1)	9.98 (2)	10.83 (4)	10.31 (3)
Death by fiery car crash**	9.78 (2)	10.60 (3)	11.00 (4)	9.77 (1)
Smashing car***	9.68 (1)	10.63 (2)	10.96 (4)	10.68 (3)
Shotgun killing*	10.25 (1)	10.42 (2)	11.41 (4)	10.46 (3)
Pistol killing***	9.93 (1)	10.78 (2)	11.46 (4)	11.23 (3)
Suffocation killing***	9.68 (1)	11.37 (3)	11.69 (4)	11.32 (2)
Sum of ranks:	(9)	(18)	(31)	(22)
		$(X^2 r = 18.75, p < .001)$		
Mean ratings across scenes:	9.44	10.36	10.96	10.53

†The larger the mean, the less acceptable the content; the higher the rank, the more acceptable.

††Mean differences for individual scenes significant by one-way analysis of variance: $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Subanalyses supported this interpretation. In seven of eight scenes the four groups were significantly different by one-way analysis of variance (see asterisks in Table 3). In each case, the variation was based on economic/race differences. The analysis of variance across the groups for the collapsed scene means in Table 3 was consistent with this result

($p<.001$). Comparing the combined lower-class subjects with the combined middle- and upper-class groups demonstrated an income difference ($p<.001$). Thus, the lower-class youngsters found the behavior in the scenes more acceptable, and this was even more true for the black disadvantaged boys.

For ordering the scenes by degrees of acceptability, the correlation among the four groups was .75 ($X^2 = 20.97$, df=7, $p<.01$). Table 3 lists the scenes in ranked order from most to least acceptable. Overall, the behavior in the furniture smashing and fist fight scenes was perceived as more acceptable; the killing scenes were least acceptable.

Professed liking. Table 4 contains the mean ratings in terms of how much the scenes were enjoyed. Here the rank order analysis of variance was marginally significant and emphasized an income difference ($X^2 = 5.74$, df=3, $p<.10$). The lower-class boys, both black and white, liked watching the violent scenes somewhat more than the middle- and upper-class boys did.

Table 4: Professed liking[†]

	Social group			
Scene[††]	Black Lower	White Lower	White Middle	White Upper
Smashing furniture	10.33 (2.5)	9.48 (1)	10.33 (2.5)	11.08 (4)
Fist fight No. 2*	9.68 (1)	10.51 (2)	11.23 (3)	12.26 (4)
Fist fight No. 1**	9.45 (1)	11.04 (3)	13.05 (4)	11.00 (2)
Smashing car	11.93 (3)	10.78 (1)	11.38 (2)	12.52 (4)
Death by fiery car crash	12.82 (2)	14.06 (4)	13.88 (3)	12.50 (1)
Shotgun killing	13.12 (1)	14.17 (4)	14.07 (3)	13.19 (2)
Pistol killing	14.30 (2)	14.02 (1)	14.81 (3)	15.10 (4)
Suffocation killing	14.98 (2)	14.51 (1)	15.38 (3)	15.48 (4)
Sum of ranks:	(14.5)	(17)	(23.5)	(25)
		($X^2r = 5.74$, $p < .10$)		
Mean ratings across scenes:	12.01	12.31	13.03	12.97

[†] The larger the mean, the less liking; the lower the rank, the less liking.

[††] Mean differences for individual scenes significant by one-way analysis of variance: *$p < .05$; **$p < .001$.

As a further check on this finding, the subanalyses were done. A one-way analysis of variance for each scene showed that two of the eight scenes were significantly different across groups (see asterisks in Table 4). This difference represented higher levels of enjoyment by the lower-class groups, principally the lower-class blacks. The eight violent scenes were collapsed and a one-way analysis of variance was computed for the group means in Table 4. This test yielded the same economic difference interpretation ($p<.05$). Third, the combined lower-class means were compared with the combined middle- and upper-class means by t-test. The more disadvantaged boys professed to enjoy the violent content more than their middle-class counterparts ($p<.01$); again, race intensified that distinction.

The scenes in Table 4 are listed in the order of enjoyment, from most to least. The correlation among the orderings of the scenes by the four groups was .88 ($X^2 = 24.50$, df=7, $p<.001$). Overall, the furniture smashing and fighting scenes were most liked; the killing scenes were least liked.

Perceived reality. Table 5 summarizes the results for this attitude dimension, which consisted of a single item and was not a strong attitudinal component as constructed. The analysis by ranks was marginally significant, but it maintained the income distinction ($p<.10$). Race did not intensify the perception differences.

Subanalyses were also less stable for this factor. For no single scene, nor for the means collapsed across scenes, was a statistically significant difference obtained. The difference between means for the combined lower classes and the combined middle and upper classes was marginally significant ($p<.10$). Overall, there was partial support for the notion that perceptions of reality vary by income level. The lower-class boys saw the violent scenes as somewhat more like real life than did the more advantaged youngsters.

Table 5 orders the scenes by amount of perceived reality, from most to least. The coefficient of concordance for similarity of order patterns across the groups was .70 ($X^2 = 19.66$, df=7, $p<.01$). Here the fist fight scenes were seen as being most real, the killing scenes among the least real.

This analysis of racial and social class differences supports these propositions: (1) black lower-class youngsters saw less violence in a given "violent" scene than did any white socioeconomic groups; (2) the lower-class boys, both white and black, saw various forms of violent behavior as more acceptable than did the middle- and upper-class white youths; (3) the lower-class youngsters enjoyed watching the "violent" scenes more; and (4) the disadvantaged boys tended to see the violent scenes as "more like real life."

There was substantial agreement among the groups about scene ordering for all four factors, exceeding .70 for the acceptability, liking, and

Table 5: Perceived reality[†]

	Social group			
Scene	Black Lower	White Lower	White Middle	White Upper
Fist fight No. 2	1.88 (1)	1.93 (2)	2.19 (4)	1.97 (3)
Fist fight No. 1	1.84 (1)	2.08 (3)	2.19 (4)	1.96 (2)
Smashing car	2.10 (3)	1.85 (1)	2.08 (2)	2.13 (4)
Shotgun killing	2.10 (2.5)	2.04 (1)	2.10 (2.5)	2.23 (4)
Pistol killing	2.40 (4)	2.17 (2)	2.33 (3)	2.16 (1)
Death by fiery car crash	2.02 (1)	2.33 (2)	2.60 (4)	2.42 (3)
Suffocation killing	2.43 (4)	2.20 (1)	2.40 (3)	2.32 (2)
Smashing furniture	2.22 (1)	2.46 (2)	2.55 (4)	2.50 (3)
Sum of ranks:	(17.5)	(14)	(26.5)	(22)
		($X^2r = 6.64, p < .10$)		
Mean ratings across scenes:	2.12	2.14	2.30	2.20

[†] The larger the mean, the less like real life; the lower the rank, the less like real life.

reality factors. This overall ordering similarity enabled examination of the minor hypothesis that weapon-induced aggression would be considered more violent than nonweapon aggression. Two tests were made from the available data. First, the violence ratings for the pistol and shotgun killing scenes were compared to the combined ratings of the two fist fight scenes. Second, all eight violent scenes were classified as either weapon or nonweapon scenes and the two sets were compared. In both comparisons, the difference was significant, as we predicted ($p < .001$).

The subhypothesis that nonweapon scenes would be seen as more real than weapon scenes was tested in the same manner. The nonweapon scenes headed the perceived reality scale and the differences were consistent ($p < .01$).

Violent vs. nonviolent scenes

The violent (experimental) scenes and the nonviolent (control) scenes were compared (by correlated t-tests) to examine the boys' relative perceptions of the two for the four viewing factors. These data are in Table

6, in terms of collapsed mean values for the violent and nonviolent scenes. The table also presents the results of a one-way analysis of variance across groups for the experimental and control scenes.

These (specific) nonviolent scenes were rated less violent, more acceptable, and better liked by each of the four social class/race groupings. That is, for each of the first three rows in Table 6, the paired means in each cell are significantly different beyond the .001 level. The reality factor differences were less stable; all but the black youngsters perceived the violent scenes as significantly less real. However, the choice of control scenes was arbitrary, and these findings of differences between the experimental and control scenes do not follow from any theoretical propositions. Although it was expected that the control scenes would be rated less violent consistently, perceptions of reality, enjoyment, etc., could vary greatly with the kind of less violent scenes used as controls.

Table 6: Mean judgments of control and experimental scenes[†]

		Black Lower	White Lower	White Middle	White Upper	F	P
Violence	C	7.38	6.55	5.77	6.27	10.15	.001
	E	11.81	13.08	13.25	12.98	9.58	.001
Acceptability	C	6.43	6.08	6.13	6.08	1.17	N.S.
	E	9.44	10.36	10.96	10.53	15.80	.001
Enjoyment	C	9.60	9.56	10.38	9.44	2.80	.05
	E	12.01	12.31	13.03	12.97	2.59	.05
Reality	C	2.04	1.90	2.07	2.05	0.95	N.S.
	E	2.12	2.14	2.30	2.20	1.03	N.S.

[†] The larger the mean the:

—more violence
—less acceptability
—less liking
—less reality

C=Control (nonviolent scenes); E=Experimental (violent scenes).

More important, the data in Table 6 enable us to examine whether the obtained social class/race differences in perception of violent content are different from those youngsters' perceptions of any other kind of television content. If the more disadvantaged youngsters saw less violence, greater acceptability, etc., in the control scenes as well as in the experimental scenes, then the stated findings could be highly artifactual. They could be totally the result of a response set to television, not predicated on the background characteristics.

The data disconfirm such a view. The means for the control scenes do not have the same pattern as those for the experimental scenes. The lower-class youths, particularly the blacks, actually judged the control scenes as more violent than the middle-class youngsters did ($p<.001$). For the other three dimensions, no consistent class or race pattern is evident for the control scenes. Therefore, this evidence suggests that the alternative explanation of a generally different response set or threshold judgment difference may be rejected.

SUMMARY

This study examined differences in perception of television violence as they related to the viewer's social class, his race, and the content of the scene.

Boys from low-income families differed from their middle-class counterparts in that the former perceived the behavior in violent scenes as more acceptable, saw violent scenes as more like real life, and liked watching the violent scenes more.

Low-income blacks differed from low-income whites in that the former saw less violence in scenes of violence.

Scenes of violence with weapons were judged more violent and less real than weaponless scenes.

In addition, the more disadvantaged youngsters liked watching all scenes—violent or not—more than their comparison group did. Yet their lesser perceptions of violence in the more violent scenes was not offset by lesser judgments of violence in the less violent scenes.

Middle-income boys saw more violence in nonweapon scenes than low-income boys did. However, since they saw more violence in general, this is a minor finding.

DISCUSSION

The data clearly support several propositions about the impact of class environment on the predispositions of these young viewers toward televised violence. Less clear is the extent to which a racial difference exists beyond the economic one. The analyses do further separate the black low-income youngsters from their white peers. (The blacks see even less violence, find it more acceptable, and like watching it more.) But an impression remains that the black respondents were a level or two down the economic scale from their white counterparts. Individual income data were not obtained, so this notion cannot be tested more directly in this study. If the blacks *were* even more economically disadvantaged, then perceptual differences reported here as extended by race may be an intensification of the class variable. The data are equivocal on that specific issue.

Earlier studies isolated age and background as important variables in the perception of violence (Toch, 1961; Reif, 1967; Moore, 1966). A recent television survey (Greenberg and Gordon, 1971) supported these experimental findings, pointing out that men perceived less violence across a set of "violent" programs than did women. That survey argued that what makes age and environment important are the related socialization experiences. For example, men learn to deal with aggression or to be aggressive differently than women. Physical aggression is a more commonplace mode of behavior for men; verbal aggression may be the balancing tool for women, but we know little about it.

Combining the variables of age, sex, and family social class into a single paradigm may illustrate better the present approach and begin to specify needed research:

A	male female	who is of	preschool teen adult	age and whose background or upbringing
has been	lower-class middle-class	which	facilitates minimizes	exposure to or experience with
violence will perceive	less more	violence in a given television scene.		

These elements may be thought of as "what the individual brings to the medium." For example, the present evidence shows that the preteen from a lower-class background, which increases exposure to violent behavior, will perceive less violence in a violent television scene. If he is black, he probably perceives even less violence. Far less is known about other combinations of these factors; very little is known about other kinds of media effects.

Recall that, in the present study, those boys who saw the violent scenes as less violent saw the nonviolent scenes as more violent; in another study, men saw more violence in nonviolent programs than did women. As a plausible explanation, we propose that if an individual's environment is hostile and frequently contains violent behavior, over time his accommodation to that setting results in his seeing a given incident of violence as less intense. Yet at the same time, toward the less violent end of the judgment continuum there is a greater than average tendency to see some violence or hostility. The experiences of the disadvantaged youngster may predispose him to see some hostility in everything, though not as much, and not with the same degree of differentiation as more general norms would indicate. Or perhaps the violence judgment dimension is just not as readily used. We argue the first explanation—that the more hostile background and environment of the youngster predisposes a more aggressive, violent outlook toward society.

Such an outlook should be reflected in the youngster's attitude toward violence or his willingness to advocate the use of violence in a given situation. Dominick and Greenberg (1970), looking specifically at attitudes toward aggression, found a relationship between the extent to which clear family norms concerning violent behavior existed and the likelihood that a child would advocate using violence in a specific situation. Where norms were lacking, the child was more favorably inclined toward aggressive behavior. The family norm related to the child's perception of the effectiveness of violence as a mode of conflict resolution; this perception in turn was correlated with his exposure to television violence. For both middle- and lower-income boys, violence was perceived to be maximally effective when television exposure was high or when family attitudes toward violence were unclear.

This study of attitudes toward violence suggests that a child's attitude toward violence will reflect his willingness to behave aggressively. Research concerning attitudes toward aggression is intermediate between determining a child's exposure to and perceptions of mediated violence and his subsequent behavior. Perceptions of violence may not coincide with attitudes toward aggression, but the researcher has little basis for suspecting otherwise.

To the extent that these linkages exist, the current research may be related to prior experimental work on aggression and imitation. Berkowitz (1962) has specified several factors which influence the likelihood that a person will behave aggressively. Once the relationships among exposure, perception, attitude, and behavior are more fully understood, they may be tied to such factors as those Berkowitz proposed. Figure 2 is a preliminary step in that direction.

Take a
male
female
who is of
preschool
teen
adult
age and whose background or upbringing has been
lower-class
middle-class
which
facilitates
minimizes
exposure to or experience with violence.

Insult
Frustrate
Don't frustrate
that person and expose him to
high
low
violence or aggression by someone
like
unlike
himself which is
justified
unjustified
and
rewarded
punished
Then, place the person in a situation which is
similar
dissimilar
to the media setting and in which
uninhibiting
inhibiting
norms are functioning. As a result, the individual will perceive
low
high
violence and will exhibit
high
low
aggressive behavior.

Figure 2: Maximizing conditions for aggressive behavior

The number of variables presented emphasizes the complexity of the issue. Surely research of a multivariable nature is required to identify the more important variables and to eliminate the lesser contributors to the behaviors examined. Some variables are manipulable, others may be unchangeable. Knowing responses to mediated violence among specifiable and major subgroups of the audience is apparently critical to the general question of effects.

Beyond this, inferences become even more speculative. If indeed, as evidence shows, the more disadvantaged are more aggressive in attitude and experience; if this aggressiveness is strongly reinforced through a steady exposure to television fare of their choosing; and if few counter-aggression messages are received from family, peers, or other socializing agencies, then the consequences are of paramount social importance.

FOOTNOTES

1. The research upon which this report is based was performed pursuant to Contract No. HSM 42-70-32 with the National Institute of Mental Health, Health Services and Mental Health Administration, U.S. Department of Health, Education and Welfare. The research staff for the project included Joseph Dominick, Liz Monroe, Marian Schaffer, Stuart Surlin, Mantha Vlahos, and Edward Wotring. We are grateful to the public schools of Grand Rapids, Kalamazoo, and Mt. Pleasant, Mich., for their cooperation on this project.
2. Each of the four scenes of violence in the first content version was tested against its counterpart scene in the second version. This was done for each of the four attitudinal dimensions. By t-tests, nine of the 16 pairs of means were significantly different. This precluded any collapsing of the two versions, and the analyses maintain this distinction.

REFERENCES

Bandura, A., and Walters, R. H. *Social learning and personality development.* New York: Holt, Rinehart and Winston, 1963.

Becker, S.L., and Wolfe, G.J. Can adults predict children's interest in a television program? In Schramm, W. (Ed.) *The impact of educational television.* Urbana: University of Illinois Press, 1960.

Berkowitz, L. *Aggression: a social-psychological model.* New York: McGraw-Hill, 1962.

Berkowitz, L. Aggressive clues in aggressive behavior and hostility catharsis. *Psychological Review*, 1964, 71, 104-22.

Buss, A.H. *The psychology of aggression.* New York: Wiley, 1961.

Carter, R.F. *The perceived appeals of television program content* (Research Bulletin No. 8). Madison: University of Wisconsin Television Laboratory, 1957.

Chilman, C.S. Child rearing and family relationship pattern of the very poor. *Welfare in Review,* 1965, **3,** 9-19.

Clark, K.B. *Dark ghetto: dilemmas of social power.* New York: Harper & Row, 1965.

Dominick, J.R., and Greenberg, B.S. Attitudes toward violence: the interaction of TV exposure, family, attitudes, and social class. In *Television and social behavior:* a report to the Surgeon General's Scientific Advisory Committee, Vol. 3. Washington, D.C.: U.S. Government Printing Office, 1971.

Epstein, S. Comments on Dr. Bandura's paper. In James, M.R. (Ed.) *Nebraska symposium on motivation.* Lincoln: University of Nebraska Press, 1962.

Flanders, J.P. A review of research on imitative behavior. *Psychological Bulletin,* 1969, 315-37.

Gans, H.J. *The urban villagers: group and class in the life of Italian Americans.* Glencoe, Ill.: Free Press, 1962.

Gerbner, G. The television world. In Baker, R.H., and Ball, S. (Eds.) *Mass media and violence:* a staff report to the National Commission on the Causes and Prevention of Violence. Washington, D.C.: U.S. Government Printing Office, 1969.

Goranson, R.E. A review of recent literature on psychological effects of media portrayals of violence. In Baker and Ball (Eds.) *Mass media and violence.*

Gordon, T.F. An exploration into television violence. *Educational Broadcasting Review,* 1969, **3,** 44-48.

Greenberg, B.S., and Dervin, B. *Use of the mass media by the urban poor.* New York: Praeger, 1970.

Greenberg, B.S. The content and context of violence in the mass media. In Baker and Ball (Eds.) *Mass media and violence.*

Greenberg, B.S. Television for children: dimensions of communicator and audience perceptions. *A-V Communication Review,* 1965, **13,** 385-96.

Greenberg, B.S. and Gordon, T.F. Perceptions of violence in TV programs: critics and the public. In *Television and social behavior* (Vol. 1).

Himmelweit, H.T., Oppenheim, A.N., and Vince, P. *Television and the child.* New York: Oxford University Press, 1958.

Kahn, M.W., and Kirk, W.E. The concepts of aggression: a review and reformation. *The Psychological Record,* 1968, **18,** 559-73.

Kaufmann, H. Definitions and methodology in the study of aggression. *Psychological Bulletin,* 1965, 64.

Moles, O.C. Child training practices among low-income families—an empirical study. *Welfare in Review*, 1965, **3**, No. 12, 1-19.

Moore, M.R. Aggression themes in a binocular rivalry situation. *Journal of Personality and Social Psychology*, 1966, **3**, No. 6.

Reif, T.F. The perception of violence as a function of historical and current behavior. Unpublished master's thesis, Michigan State University, 1967.

Sears, R.R. Relation of early socialization experiences to aggression in middle childhood. *Journal of Abnormal and Social Psychology*, 1961, **63**, 466-92.

Siegel, S. *Non-parametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

Toch, H.H., and Schulte, W.H. Readiness to preceive violence as a result of police training. *British Journal of Psychology*, 1961, **52**.

Toch, H.H. The stereoscope: a new frontier in psychological research. *Research Newsletter*, 1961, **3**, Nos. 3-4.

U.S. National Advisory Commission on Civil Disorders. *Report of the National Advisory Commission on Civil Disorders*. New York: Bantam, 1968.

Weiss, W. Effects of mass media of communication. In Lindzey, G., and Aronsen, E. (Eds.) *The handbook of social psychology*. (Vol. 5). Reading, Mass.: Addison-Wesley, 1969.

Appendix A: Taped programs

DAY	TIME	NET	SHOW	DURATION
Sunday	7:00	CBS	*Lassie*	1/2hr.
"	7:00	ABC	*Land of the Giants*	1 hr.
"	8:00	ABC	*FBI*	1 hr.
"	9:00	NBC	*Bonanza*	1 hr.
"	10:00	NBC	*The Bold Ones*	1 hr.
"	10:00	CBS	*Mission: Impossible*	1 hr.
Monday	7:30	CBS	*Gunsmoke*	1 hr.
"	7:30	ABC	*It takes a Thief*	1 hr.
Tuesday	7:30	CBS	*Lancer*	1 hr.
"	7:30	ABC	*Mod Squad*	1 hr.
"	10:00	ABC	*Marcus Welby*	1 hr.
Wednesday	7:30	NBC	*Virginian*	1 1/2 hrs.
"	9:00	CBS	*Medical Center*	1 hr.
"	10:00	CBS	*Hawaii Five-0*	1 hr.
"	10:00	NBC	*Then Came Bronson*	1 hr.
Thursday	7:30	NBC	*Daniel Boone*	1 hr.
"	8:30	NBC	*Ironsides*	1 hr.
"	9:30	NBC	*Dragnet*	1/2 hr.
"	10:00	ABC	*Paris 7000*	1 hr.
Friday	7:30	CBS	*Get Smart*	1/2 hr.
"	8:30	NBC	*Name of the Game*	1 1/2 hrs.
"	10:00	NBC	*Bracken's World*	1 hr.
Saturday	8:30	NBC	*Adam-12*	1/2 hr.
"	10:00	CBS	*Mannix*	1 hr.
			Total time	23 hrs.

Appendix B: Scene descriptions

First Scene (Practice—same for both versions): A girl is being held captive by two men in a remote cabin. The girl breaks free and runs into the woods. The men chase her from different directions. Rapid cutting builds an air of suspense until one man, laughing, jumps from behind some bushes and grabs the startled girl, face-to-face in a bear hug. (Taken from *The FBI*—1 min. 5 sec.)

Second Scene (Version One): In a plush business office an angry man in a business suit confronts two other business men. The angry man begins shouting and smashing furniture with his bare hands, as the other men look on in dismay. (Laugh track deleted from test scene. Taken from *Get Smart*—21 sec.)

Second Scene (Version Two): In the early morning quiet of a city street, a woman in robe and with hair in curlers approaches a car parked near a bar. Shouting about her no-good drunkard husband she begins smashing the glass and fenders of the car with a baseball bat. A police car approaches, two policemen get out and subdue the woman. (Taken from *Adam-12*—25 sec.)

Third Scene (Control—same for both versions): A boy and a dog walk slowly past some adobe houses. Peaceful music accompanies them as they stroll into a wooded area in the shadows of late afternoon. (Taken from *Lassie*—25 sec.)

Fourth Scene (Version One): In the large stately house, a man glares at a group of his peers. In admission of his guilt he screams, "All right, I did it, I killed her." A friend tries to stop him as he runs from the room and is knocked to the floor. Running from the room, he pushes a button to open the huge iron gates to the manor, jumps in his car and speeds off. Through a malfunction, the gates fail to open and the car crashes into the gates and bursts into flames. (Taken from *Name of the Game*—37 sec.)

Fourth Scene (Version Two): A burglar, in the bedroom of a sleeping young woman, is trying to remove a photograph from a glass frame. The frame slips and crashes to the floor, wakeing the woman. The burglar takes a pillow and forces it over the woman's face. With her limp body in his arms, he walks to the third-story bedroom window and drops her out. (Taken from *Paris 7000*—23 sec.)

Fifth Scene (Version One): A man with sawed-off shotgun cautiously peers around the corner in a corridor. Satisfied, he steps out, takes careful aim and pulls the trigger. Inside a glass-walled office a man is sitting behind a desk with his back to the assassin. The blast hurls the man, flying glass and debris across the desk. He ends up sprawled on the floor. (Taken from *Hawaii Five-O*—15 sec.)

Fifth Scene (Version Two): In a crowded parking lot, a man is preparing to drive away from a social gathering. Guests are standing on a nearby porch with drinks in hand. As he approaches a gate, a car pulls through the gate and stops, blocking the exit. Annoyed, he honks his horn and hollers at the guy to "move it." The second man gets out, walks around his car, pushes a pistol in the first man's face and pulls the trigger. The guests' heads turn in slow motion to the roll of a harp. (Taken from *The Bold Ones*—20 sec.)

Sixth Scene (Control—same for both versions): As a lone motorcycle rider travels down a dirt road other riders, dressed as Indians and howling, race up the road embankment as if to attack. Instead, all riders continue down the road and out of sight around a bend. (Taken from *Bronson*—22 sec.)

Seventh Scene (Version One): A man opens the door of his female companion's apartment and escorts her inside. As he turns to close the door a second man hits him on the head, knocking him dazed to the floor. The intruder grabs the girl and she struggles to get free. Regaining his senses, the woman's companion jumps on the intruder and a fist fight starts. In attempting to escape, the intruder's path lies along a long scatter rug which his pursuer pulls. Losing his footing, the intruder crashes to the floor, striking his head, and is unconscious. (Taken from *Paris 7000*—37 sec.)

Seventh Scene (Version Two): In a stylish middle-class apartment, the private eye holds a gun on the villain as he questions him. The villain relates that the action will take place at a specific hour. As the private eye glances at his watch the villain knocks the gun away and a fist fight starts. Crashing over the furniture, the lamps are knocked out and the fight continues in semidarkness. The private eye hits the villain into a semiconscious state, grabs the gun, and holds him at bay. (Taken from *Mannix*—35 sec.)

Children's Perceptions of Television Violence: a Replication

Bradley S. Greenberg and Thomas F. Gordon

Michigan State University

We have reported the results of a study designed to determine the perceptions of televised violence among preteen males from varying racial and socioeconomic backgrounds (Greenberg and Gordon, 1971). The basic rationale of that study posited that greater exposure to "real life" aggression—a more common phenomenon among the disadvantaged—manifests itself in greater tolerance for aggressive behavior, whether real or mediated.

There is much evidence to indicate that physical aggression is more familiar to youngsters from low-income environments. In such homes, physical punishment is used to control behavior more often than is a verbal approach; outside the home, the environment is more likely to be hostile for the low-income child (Chilman, 1965; Clark, 1965; Gans, 1962; Moles, 1965; Sears, 1961; National Advisory Commission on Civil Disorders, 1968).

The youngster who has been exposed more frequently to greater amounts of real-life aggression may have a higher tolerance for television violence. For these reasons, the earlier study and the present one[1] tested these hypotheses:

H1: The less advantaged youngster will perceive less violence in a given segment of television violence than will a middle-class youngster.

H2: The less advantaged youngster will judge mediated violence as a more acceptable mode of behavior than the middle-class viewer.

H3: The less advantaged youngster will see television violence as being more realistic than his middle-class counterpart.

It was thought that the disadvantaged youngster should be particularly attracted to television content which is high in action and excitement. Programs which contain heavy doses of violence should be more arousing, given a home environment in which passivity and verbal interaction are secondary to direct action. These hypotheses were examined:

H4: The less advantaged youngster will judge violence as more enjoyable to watch.

H5: The less advantaged youngster will testify to more self-arousal from television violence.

Separate hypotheses for racial differences were not stipulated. It was assumed that, among the poor, blacks are even more disadvantaged. Thus, they should exhibit a higher degree of the postulated behaviors.

Finally, we tested two hypotheses extracted from Himmelweit (1958) and Berkowitz (1964):

H6: Violent scenes with weapons will be perceived as more violent than violent scenes without weapons.

H7: Violent scenes with weapons will be perceived as less realistic than violent scenes without weapons.

In the first study, conducted with 325 fifth grade boys, these major results were found for the hypotheses:

H1: No social class difference, but a sharp racial distinction, with the low-income black youngsters perceiving significantly less violence in the same scenes than any other comparison group.

H2,3,4: The more disadvantaged youngsters, whether white or black, judged the behavior in violent scenes as more acceptable, saw violent scenes as more like real life, and liked watching the violent scenes more.

H5: No arousal factor emerged from the attitudinal data.

H6,7: Scenes of violence with weapons were judged as more violent and less real than weaponless scenes.

The present study encompasses the same hypotheses, applied to a different population group. We shifted upward in age to determine whether results obtained with ten-year-olds would be replicated with a teenage group, specifically 14-year-olds. The basic purpose of the study was to examine the generalizability of the earlier results.

METHODS

The methods were identical to those implemented earlier with the fifth-grade boys (see our *Social Class and Racial Differences*, in this volume). In general, boys in their early teens were shown a sequence of television vignettes composed of violent and nonviolent scenes. The experimental scenes ranged from 15 to 37 seconds each. Viewing was done in groups of six to eight boys in a room in their public school. The youngsters' responses were in terms of verbal scale ratings of each of the scenes. The schools from which the youths came (and hence the youths themselves) were differentiated in terms of socioeconomic status and race.

Video materials

Two versions of videotaped materials were used. Each tape contained an identical practice stimulus and two identical control (nonviolent) scenes. Each version had four different scenes of violence, matched for content and length. The kinds of violence depicted in each version included destruction of property, physical assault against others, and intentional killing of others. All scenes came from commercial prime time television programs. The order of presentation for each version, held constant, was: practice scene, one violent scene, control scene, two violent scenes, control scene, and final violent scene.

Subjects

The public schools in Kalamazoo, Michigan, provided the subject pool. A total of 263 eighth grade boys were used; they attended the three junior high schools in the community designated by the superintendent's office as primarily containing pupils from lower, middle, and upper-middle income families. Parental nonpermission slips were used; but only a trivial number of nonpermissions were received. By race and social class, the subjects were distributed as follows: 66 black lower class; 78 white lower class; 37 white middle class; and 82 white upper-middle class.

Instrument

The test instrument was identical to that used in the earlier study. It consisted of 15 items, three each for five hypothesized attitudinal dimensions. The following are sample items for each dimension:

Degree of violence
Was what you saw ____ Extremely violent
____ Very violent
____ Pretty violent
____ Not very violent

Acceptability of the behavior
Is it ____ Very right for people to be this way
____ A little right
____ Not very right
____ Not right at all for people to be this way

Enjoyment of the content
Was what you saw
a show like ____ You really like to see
____ You sometimes like to see
____ You don't like to see very much
____ You don't like to see at all

Perceived reality
What you saw was ____ Not very exciting
____ A little exciting
____ Very exciting
____ Extremely exciting

Degree of arousal
What you saw was ____ Very much like real life
____ Pretty much like real life
____ Not much like real life
____ Not at all like real life

The complete set of items is in Table 1, together with empirical verification of the dimensions from both this and the prior study.

Procedures

Testing was conducted in January 1971. Each school provided a room large enough for the videotape equipment and for six to eight boys seated in front of a television set. The boys were told that we wanted their reactions to scenes from regular television programs, that this was not a test and would in no way affect their classroom evaluation. They made no personal identification on the instrument. Booklets were coded for race and version after the boys left the viewing room.

Subjects first viewed the practice scene. The experimenter completed two or three items with them to clarify how they were to proceed. The boys completed the remaining items for the practice scene and were questioned about difficulties with words or procedures. The subjects were then shown the six remaining scenes and rated each scene immediately after viewing it. On the average, it took 25-30 minutes to view and rate all seven scenes. After completing the ratings, the boys were asked not to talk to their classmates about what they had done until everyone had participated. Teachers were asked not to discuss the children's experiences with the class until testing in that school was completed.

RESULTS

Four major analyses were completed: (1) a factor analysis of the test items; (2) a comparison of racial and social class differences in response to the stimuli; (3) an examination of differences among kinds of violence; and (4) a check on relative perceptions of the control and experimental scenes. Results are presented in that order. Each is compared to the results of the first study.

Item analysis

Responses to all items, across all respondents, for all violent scenes, were intercorrelated and then submitted to a principal axis factor analysis with varimax rotation. This was done to determine the extent to which the attitude items were used by the respondents in the same way they were conceived by the investigators. A summary of this analysis is in Table 1.

In the first study, there were three major factors and one minor one. In this replication, there were three major factors and two minor ones.

A principal attitude dimension was that of perceived *violence.* In each study, the same four violence items loaded primarily on this factor. These items tapped the judged anger, violence, cruelty, and seriousness apparent to the respondents. The first three items were designated, *a priori,* as a violence perception factor. This factor accounted for about one-sixth of the total variability in judgment of the television stimuli.

A second major factor was that of judged *acceptability* of the content. In both the original study and this replication, the three items originally designed to tap this area of judgment fit best on this single factor. Together, they accounted for another one-sixth of the total variance.

In the original study, the third factor was professed *enjoyment* of the content depicted in the scenes. Five scales formed the original factor; in this study, three of the same five items persisted and are considered a liking dimension. They are the three items designed to tap this dimension of attitude, and they account for another one-sixth of the respondents' judgmental variation.

Table 1: Factor items

Factor 1. Perceived violence		Factor loadings	
ITEMS:		Study 1	This Study
Were the people	___ Not very angry ___ A little angry ___ Very angry ___ Extremely angry	.64	.69
Was what you saw	___ Not very violent ___ Pretty violent ___ Very violent ___ Extremely violent	.58	.84
Was what you saw	___ Not very serious ___ A little serious ___ Pretty serious ___ Very serious	.71	.73
Was what you saw	___ Not very cruel ___ A little cruel ___ Pretty cruel ___ Very cruel	.58	.72
	% Total Variance	(17%)	(18%)
Factor 2. Perceived acceptability		**Factor loadings**	
ITEMS:		Study 1	This Study
Is it	___ Very right for people to be this way ___ A little right ___ Not very right ___ Not right at all for people to be this way	.84	.78

Was what you saw	___A very good thing to do ___A pretty good thing ___A pretty bad thing ___A very bad thing to do	82	.82
Is it	___Very nice for people to act like this ___Pretty nice ___Not very nice ___Not nice at all for people to act like this	81	.81
	% Total Variance	(17%)	(18%)

Factor 3. Professed liking . . . . Factor loadings

ITEMS:		Study 1	This Study
What you saw was	___A very good thing to watch ___A pretty good thing ___A pretty bad thing ___A very bad thing to watch	77	.75
Was it	___A wonderful show ___A pretty good show ___A pretty bad show ___A terrible show	80	.85
Was what you saw a show like	___You really like to see ___You sometimes like to see ___You don't like to see very much ___You don't like to see at all	79	.86
	% Total Variance	(*)	(17%)

Table 1: Factor items (Continued)

Factor 4. Perceived humor		Factor loadings	
ITEMS:		Study 1	This Study
What you saw was	____A very funny thing to see ____A pretty funny thing ____A pretty sad thing ____A very sad thing to see	.77	.85
Does what you saw	____Make you feel like laughing a lot ____Make you feel like laughing a little ____Not make you feel like laughing very much ____Not make you feel like laughing at all	.70	.88
	% Total Variance	(*)	(13%)

Factor 5. Perceived reality		Factor loadings	
ITEM:		Study 1	This Study
What you saw was	____Very much like real life ____Pretty much like real life ____Not much like real life ____Not at all like real life	.87	.86
	% Total Variance	(7%)	(7%)

The following items were too impure to assign to a single factor:

What you saw was	____Not very exciting ____A little exciting ____Very exciting ____Extremely exciting
Was what you saw	____Very much for fun ____Pretty much for fun ____Not very much for fun ____Not for fun at all

---

*Factor 4 items were both in Factor 3 in Study 1. The 5 items together accounted for 25 percent of the total variance.

The two items which dropped off the enjoyment factor of the first study loaded as a fourth factor. Both scales—how funny the scenes were and how much they made the respondent feel like laughing—were originally constructed as part of an arousal index. Given the humor component in each of them, it appears appropriate to relabel this two-item factor as perceived *humor.*

The final factor was but a single item—as it had been in the first study. It assessed the *reality* of the television scenes.

Two items were impure (equally in both studies), so they were dropped from all subsequent analyses. No interpretable arousal factor appeared in either study.

In the original study, the four emergent factors—violence, acceptability, enjoyment, and reality—accounted for 66 percent of the total variance. In the present study, the five factor solution (the fifth being humor) accounted for 73 percent of the total variance in judgment.

Social class and racial differences in perceptions of televised violence

The basic proposition of this study was that attitudinal assessments of television violence would order among the male respondent groups in this fashion: upper-income whites, middle-income whites, lower-income whites, lower-income blacks. Overall class differences were predicted, with the racial comparison expected to intensify such differences.

Item scores were summed for those items on each of the five dimensions of judgment extracted through the factor analytic procedure. To analyze the two versions of the experimental stimuli, given repeated measures within each version, a Friedman two-way rank-order analysis of variance was used (Siegel, 1956). After this basic analysis, we also did subanalyses for each of the five dependent variables which: (1) compared the respondent groups within each scene; (2) compared the respondent groups by collapsing across the eight replicate scenes; and (3) compared the two combined lower-income groups with the two combined higher-income groups.

Perceived violence. Table 2 presents the mean values for the perceived violence dimension. The data do not support the hypothesis for this 14-year-old test group. Less violence was not perceived by the lower-income groups. This is the only major portion of the findings which fails to replicate from the earlier study. The four groups do not order either in a linear or in a curvilinear fashion. If anything, the boys from the highest income group reported the violent scenes as less violent. (In the earlier study, it was the blacks who reported least violence, with the three white groups roughly comparable in their estimates.) For these respondents, no single scene produced significant differences in perceived violence. The means collapsed across the eight scenes did not yield differences. The social class comparison—the two lower-class groups *vs.* the two higher-class groups—did not yield differences.

Table 2: Perceived violence[†]

	Social group			
Scene	Black Lower	White Lower	White Middle	White Upper
Shotgun killing	$13.08^{(3)}$	$12.80^{(1)}$	$13.32^{(4)}$	$13.06^{(2)}$
Pistol killing	$12.93^{(3)}$	$12.79^{(1)}$	$13.93^{(4)}$	$12.80^{(2)}$
Death by fiery car crash	$12.68^{(2)}$	$13.14^{(3)}$	$13.32^{(4)}$	$12.55^{(1)}$
Fist fight No. 1	$13.08^{(4)}$	$12.71^{(2)}$	$12.86^{(3)}$	$12.55^{(1)}$
Suffocation killing	$12.72^{(3)}$	$12.05^{(1)}$	$13.53^{(4)}$	$12.31^{(2)}$
Smashing car	$12.31^{(3)}$	$12.28^{(2)}$	$12.40^{(4)}$	$11.66^{(1)}$
Smashing furniture	$12.03^{(3)}$	$12.51^{(4)}$	$11.13^{(1)}$	$11.32^{(2)}$
Fist fight No. 2	$11.07^{(1)}$	$11.72^{(3)}$	$11.87^{(4)}$	$11.23^{(2)}$
Sum of ranks:	(22)	(17)	(28)	(13)
		($X_r^2 = 9.45, p < .05$)		
Mean ratings across scenes:	12.52	12.47	12.77	12.21

[†] The larger the mean, the more violence; a rank of 1 equals least perceived violence.

The scenes in Table 2 are ordered from most to least violent. The order correlates .83 with that from the first study. The two weapon murders were perceived as most violent, a fist fight and a furniture-breaking scene as least violent. The mean rating of violence in these scenes for the four respondent groups was approximately 12.50, on a scale ranging from 4-16. It was approximately 13.0 for the earlier study, indicating somewhat less overall violence judged by the older boys. Thus, among the eighth graders, judgments of amount of violence for a set of television vignettes did not relate to the race or income background of the viewers in any meaningful fashion.

Perceived acceptability. Table 3 presents the data for this attitude factor. The analysis of variance by ranks was significant across the four groups for the eight violent scenes ($p < .01$). This difference was solely one of race. The black youngsters found the scenes significantly more acceptable than did any other viewer group.

In the earlier study, there was a parallel race difference. As the ranks and means in Table 3 indicate, a racial difference existed for seven of eight scenes. Although the differences for any one scene were not statistically significant, the collapsed means across the eight scenes corroborated the racial difference ($p < .10$).

The scenes are ordered in Table 3 from most to least acceptable, and the order correlates .98 with the scene ordering in the fifth graders' study. Two scenes switched adjacent order positions. Least acceptable

Table 3: Perceived acceptability†

	Social group			
Scene	Black Lower	White Lower	White Middle	White Upper
Fist fight No. 2	8.17$^{(1)}$	8.67$^{(4)}$	8.47$^{(3)}$	8.34$^{(2)}$
Smashing furniture	9.11$^{(2)}$	9.80$^{(4)}$	9.09$^{(1)}$	9.45$^{(3)}$
Fist fight No. 1	9.24$^{(1)}$	9.71$^{(3)}$	9.64$^{(2)}$	10.04$^{(4)}$
Smashing car	9.55$^{(1)}$	10.28$^{(4)}$	9.87$^{(3)}$	9.66$^{(2)}$
Death by fiery car crash	9.38$^{(1)}$	10.20$^{(3)}$	10.23$^{(4)}$	10.02$^{(2)}$
Shotgun killing	10.22$^{(1)}$	10.77$^{(2)}$	10.82$^{(3)}$	10.91$^{(4)}$
Pistol killing	10.55$^{(1)}$	10.98$^{(3)}$	11.13$^{(4)}$	10.80$^{(2)}$
Suffocation killing	10.62$^{(1)}$	11.14$^{(3)}$	11.33$^{(4)}$	11.09$^{(2)}$
Sum of ranks:	(9)	(26)	(24)	(21)
		($X_r^2 = 13.05, p < .01$)		
Mean ratings across scenes:	9.59	10.20	10.05	10.05

†The larger the mean, the less acceptable the content; a rank of 1 equals most acceptable.

were the scenes of killings, in maximum contrast to the fighting scenes and furniture breaking.

The scale of acceptability had a range of 3-12. The mean for the black youths in the present study was 9.59; it was 9.44 in the earlier study. The means for the three other groups exceeded 10.0 in each study. No group judged the behavior as acceptable, but the black youngsters were less negative in all comparisons.

Professed liking. Table 4 contains the mean ratings in terms of how much the scenes were enjoyed. Here, the rank order analysis of variance was marginally significant and emphasized a racial difference ($p < .10$). The highest level of liking for these violent scenes came from the black youths. Second, the eight violent scenes were collapsed and a one-way analysis of variance was computed for the group means in Table 4. The significant difference ($p < .001$) provided stronger support for the posited interpretation of the differences.

For four of the individual scenes, the set of means is consistently different, with the black youths on one extreme and the upper-income white youths on the other. If one collapses the data for the two lower-income groups and compares them with the collapsed middle and upper groups, the difference between income groups is significant ($p < .02$). However, the main contributor to this difference remains the striking divergence between the blacks and the upper-income whites. In the first study, the difference was more one of income than of race in professed liking of the content. Here, it is primarily a difference between the extreme income groups.

Table 4: Professed liking†

	Social group			
Scene ‡	Black Lower	White Lower	White Middle	White Upper
Fist fight No. 2	$4.17^{(1)}$	$4.98^{(4)}$	$4.87^{(3)}$	$4.71^{(2)}$
Fist fight No. 1*	$4.27^{(1)}$	$5.46^{(2)}$	$5.64^{(3)}$	$5.72^{(4)}$
Smashing car	$5.21^{(1)}$	$5.63^{(3)}$	$5.47^{(2)}$	$5.83^{(4)}$
Smashing furniture*	$5.08^{(1)}$	$6.43^{(3.5)}$	$6.23^{(2)}$	$6.43^{(3.5)}$
Death by fiery car crash*	$4.92^{(1)}$	$6.43^{(3)}$	$5.91^{(2)}$	$6.66^{(4)}$
Shotgun killing*	$5.19^{(1)}$	$6.77^{(2.5)}$	$6.77^{(2.5)}$	$6.89^{(4)}$
Pistol killing	$7.07^{(4)}$	$6.42^{(1)}$	$6.67^{(2)}$	$6.86^{(3)}$
Suffocation killing	$7.00^{(4)}$	$6.81^{(3)}$	$6.67^{(1)}$	$6.71^{(2)}$
Sum of ranks:	(14)	(22)	(17.5)	(26.5)
		($X_r^2 = 6.64$, $p < .10$)		
Mean ratings across scenes:	5.30	6.10	6.05	6.26

† The larger the mean, the less liking; a rank of 1 equals most liking.

‡ Mean differences for individual scenes significant by one-way analysis of variance: *$p < .003$.

The scenes are listed in Table 4 in order from most enjoyed to least enjoyed. This order correlates .85 with the first study. Most liked were the fight scenes; least liked were the killings.

In terms of the level of liking for the content, the scale range was 3-12. The black youngsters were saying that the scenes were between "very good" and "pretty good" things to watch. The white viewers were saying the scenes were between "pretty good" and "pretty bad" to watch. The means for all the eighth grade groups reflected more enjoyment of this kind of content than the means obtained for the parallel fifth grade groups.

Perceived humor. Table 5 contains the results of this analysis. (The reader will recall that this factor exists only for the present study. The two items which form this factor were part of the enjoyment factor in the original study.)

The basic analysis indicates a clear-cut income difference in the perceived humor of the violent scenes; the youngsters from more disadvantaged homes perceived significantly more humor in the scenes (p. $< .02$). For four individual scenes, this pattern approaches significance; across the eight scenes, the combined lower-class youths perceive significantly more humor ($p < .02$). Thus, the humor results parallel the perceived enjoyment results of the original study.

The scenes are ordered in Table 5 from most perceived humor to least. The funniest scene, according to the respondents, was the depiction of a woman smashing a car with a baseball bat; least humorous were the killings.

Table 5: Perceived humor[†]

	Social group			
Scene	Black Lower	White Lower	White Middle	White Upper
Smashing car	$3.38^{(1)}$	$3.72^{(2)}$	$3.87^{(4)}$	$3.83^{(3)}$
Smashing furniture	$3.38^{(1)}$	$3.97^{(4)}$	$3.91^{(3)}$	$3.74^{(2)}$
Fist fight No. 1	$4.76^{(1)}$	$4.77^{(2)}$	$5.82^{(4)}$	$5.15^{(3)}$
Fist fight No. 2	$4.59^{(1)}$	$4.93^{(2)}$	$5.60^{(4)}$	$5.40^{(3)}$
Death by fiery car crash	$5.57^{(1)}$	$6.20^{(2)}$	$6.32^{(3)}$	$6.38^{(4)}$
Shotgun killing	$5.89^{(1)}$	$6.31^{(2)}$	$6.82^{(4)}$	$6.55^{(3)}$
Pistol killing	$6.62^{(3)}$	$5.86^{(1)}$	$6.73^{(4)}$	$6.49^{(2)}$
Suffocation killing	$7.03^{(4)}$	$6.00^{(1)}$	$6.27^{(2)}$	$6.51^{(3)}$
Sum of ranks:	(13)	(16)	(28)	(23)
		($X_r^2 = 10.35$, $p < .02$)		
Mean ratings across scenes:	5.12	5.21	5.68	5.50

[†] The larger the mean, the less humor; a rank of 1 equals most humor.

On the question of how much humor is seen in violence, the scale range was 2-8. The two disadvantaged groups were at the midpoint of the scale, with the remaining groups about one-half unit on the "unfunny" side of the scale.

Perceived reality. Table 6 summarizes the findings. Although a single item comprises this attitude, the findings sharply differentiate perceptions of reality by both income and race. The black youngsters are more prone to say that what they saw was a close approximation of real life than are the white youngsters. The highest income grouping of white youngsters were least likely to perceive the violent scenes as realistic, less likely than their white low-income counterparts ($p < .01$).

For six of eight individual scenes, one-way analyses of variance were significant with the same general pattern. Collapsing across all eight scenes yielded the same ordering ($p < .001$). The final subanalysis, comparing the two lower-income groups with the two higher groups, was also consistent with the above findings ($p < .01$).

The scenes are listed in Table 6 in order from most to least realistic. Maximum reality was judged to exist in fist fights and car-smashing vignettes, least in the furniture breaking and car crash. The order of the scenes correlated .90 with the ordering obtained on this dimension from the fifth-grade boys.

The reality item was, "What you saw was (very much, pretty much, not much, not at all) like real life." This item had a scale of 1-4; the absolute scale positions of the groups placed the black youngsters as saying "pretty much like real life," and the white boys as saying "not much

Table 6: Perceived reality†

	Social group			
Scene‡	Black Lower	White Lower	White Middle	White Upper
Fist fight No. 2**	$1.62^{(2)}$	$1.93^{(3)}$	$1.60^{(1)}$	$2.29^{(4)}$
Smashing car	$1.93^{(2)}$	$1.98^{(3)}$	$1.80^{(1)}$	$2.17^{(4)}$
Fist fight No. 1**	$1.73^{(1)}$	$1.80^{(2)}$	$2.41^{(4)}$	$2.26^{(3)}$
Shotgun killing*	$1.97^{(2)}$	$1.94^{(1)}$	$2.31^{(3)}$	$2.36^{(4)}$
Suffocation killing	$1.97^{(2)}$	$2.23^{(3)}$	$1.87^{(1)}$	$2.34^{(4)}$
Pistol killing*	$2.03^{(1)}$	$2.35^{(3)}$	$2.13^{(2)}$	$2.63^{(4)}$
Death by fiery car crash**	$1.91^{(1)}$	$2.29^{(2)}$	$2.64^{(3)}$	$2.66^{(4)}$
Smashing furniture**	$2.14^{(1)}$	$2.54^{(3)}$	$2.50^{(2)}$	$2.98^{(4)}$
Sum of ranks:	(12)	(20)	(17)	(31)
		($X_r^2 = 14.55, p < .01$)		
Mean ratings across scenes:	1.92	2.13	2.22	2.48

†The larger the mean, the less like real life; a rank of 1 equals most perceived reality.
‡Mean differences for individual scenes significant by one-way analysis of variance: $^{*}p < .10$, $^{**}p < .01$

like real life." These reality levels are similar to those found in the original study.

Summary of social class and racial differences

Figure 1 highlights the comparative findings of the two studies, the first with fifth graders and the present one with eighth graders, all boys.

Dimensions of Judgment	Study 1	Study 2
1. Perceived violence	Racial difference	No difference
2. Perceived acceptability	Income and race difference	Race difference
3. Professed enjoyment	Income and race difference	Income and race difference
4. Perceived humor	(not assessed)	Income difference
5. Perceived reality	Income difference	Income and race difference

Figure 1: Summary of findings

Weapons vs. no weapons

It was posited that weapon-affiliated aggression would be considered more violent than nonweapon aggression. Two tests were made of this hypothesis. First, the violence ratings for the pistol and shotgun scenes were compared with comparable data from the two fist fight scenes. Second, all weapon-bearing scenes were compared with nonweapon scenes. In both comparisons, the differences were significant, as predicted ($p < .01$). The same finding was obtained in the first study.

It was further posited that scenes without weapons would be perceived as more realistic. This was not supported in the present study but was obtained in the earlier one.

Scene ordering across dimensions

We also examined the extent to which the ordering of scenes, as reported above for each dimension, related to the ordering of kinds of violence across dimensions. Given but eight scenes in this analysis, the magnitude of the relationships obtained are of some interest.

In both studies: (a) the ordering of scenes on the *violence* dimension was significantly negatively correlated with their ordering on the *acceptability* dimension (Rho = −.88 in Study 1 and −.69 in Study 2); and (b) the ordering of scenes on the *acceptability* dimension was significantly positively correlated with their ordering on the *enjoyment* dimension (Rho = .95 in Study 1 and .90 in Study 2).

In Study 1, the ordering of scenes on the *violence* dimension was negatively correlated with the *enjoyment* dimension (Rho = −.93). In this replication, it approaches significance (Rho = −.54). Further, in this second study, the ordering of scenes on the perceived *humor* dimension was significantly positively correlated with the ordering on both the *acceptability* (Rho = .76) and *enjoyment* (Rho = .73) dimensions and approached significance on the *violence* factor in a negative relationship (Rho = −.57).

Violent vs. nonviolent scenes

Means for each factor for the violent (experimental) scenes and the nonviolent (control) scenes are in Table 7. Each of the respondent groups found the experimental scenes significantly more violent and significantly less acceptable than the control scenes. The pattern is the same for the two upper-income groups in terms of the enjoyment dimension. The direction is the same for the two lower-income groups on that dimension, though not significant. The three groups of white respondents found the experimental scenes less humorous than the control scenes; the black low-income youngsters rated the violent and nonviolent scenes as equally humorous. There was no difference in perceived reality between the control and experimental scenes. These differences are presented for descriptive information. Only the differences on the dimension of violence were essential and followed from the logic of this study. Differences in perceived enjoyment, acceptability, and so on would have varied had other types of control scenes been used. In the present study, the control scenes consisted of a boy walking with his dog along a mountainside and a motorcyclist racing around the hills with other cyclists following him.

Table 7: Mean judgments of control and experimental scenes†

		Black Lower	White Lower	White Middle	White Upper	F	P
Violence	C	5.73	5.60	5.50	5.42	0.53	N.S.
	E	12.52	12.47	12.77	12.21	0.65	N.S.
Acceptability	C	5.92	6.14	6.22	6.01	0.67	N.S.
	E	9.59	10.20	10.05	10.05	2.22	N.S.
Enjoyment	C	5.67	6.22	6.91	7.09	8.72	.001
	E	5.30	6.10	6.05	6.26	5.70	.001
Humor	C	5.21	4.97	5.14	5.30	1.22	N.S.
	E	5.12	5.21	5.68	5.50	2.50	.10
Reality	C	1.79	2.02	2.12	2.26	6.68	.001
	E	1.92	2.13	2.22	2.48	9.34	.001

†The larger the mean the:
- more violence
- less acceptability
- less enjoyment
- less humor
- less reality

C = Control (nonviolent scenes); E = Experimental (violent scenes).

It is important to note some of the row differences obtained. The lower-class youngsters, as accentuated by the data for the blacks, liked the control scenes more than did the upper-income youngsters. For this group, it is not just the violent content that is liked more and found more realistic; it is any kind of television content. However, in terms of Acceptability, Perceived Humor, and Reality, the tendency is for the race/income differences to be particularly apparent for the experimental or violent scenes. Thus, as in the prior study, there is some evidence that the perception differences are not threshhold differences for all of television content, but hold true particularly for violent program content.

DISCUSSION

How do youngsters perceive television violence? Do perceptions differ as a function of the child's background?

The two studies show that violence is clearly recognized and labeled as violence by all groups, that the behavior is not considered acceptable, but that it is enjoyable to watch. Further, it borders on being called real-to-life.

Such perceptions, however, do differ between racial groups and between children of different income levels. Among younger, preteen black children, there are lesser perceptions of violence. Among both ten- and 14-year-olds, there is greater liking for violence and greater perceived reality in television violence among the more disadvantaged. The behavior exhibited in violent television scenes is more acceptable to the more disadvantaged. It is also considered to be more humorous.

At best, we have identified both race and income level as differentiating characteristics in such perceptions, without indicating the significance of one *vs.* the other. The most plausible conclusion we can suggest is that race intensifies the differences which would exist as a function of low income alone.

The critical question which remains is whether differing perceptions manifest themselves in differential overt behavior, particularly aggressive behavior. Subsequent studies in this area must, it seems to us, include at least two kinds of experimental efforts. One would be using natural television content stimuli (exemplified by the type used in the present study), obtaining parallel attitudinal information, and then providing an opportunity for the viewers to aggress, preferably in an antisocial fashion. In that manner, one could correlate the acceptability of the message content, the liking for it, or the perception of how violent it is, with the consequent behavior.

The second approach would be using the same type of content, but experimentally inducing the appropriate perceptions. The television scenes could be identified for the viewer as acceptable or unacceptable behavior, violent behavior, etc., and subsequent aggressive responses could be assessed. This has a theoretic linkage to the notions of forewarning as advanced by McGuire (1966). It could also be used to simulate the kind of information which parents might transmit to their children as they watch such programming.

For example, Hicks (1968) had children watch an aggressive television model in the presence of an adult. While watching, the adult made comments about the actions taking place. The comments were positive for one set of children and negative for the other. The adult remained silent in the control group. In a postviewing situation, the group hearing positive comments was more aggressive than the negative comments or control groups. The control group was also more aggressive than the negative comments group.

The Berkowitz research on justification of aggression bears some similarities to this approach. By labeling a televised act of violence as justified aggression (which we would associate with the notion of acceptability in the present study), he has consistently obtained more aggressive responses than when the same act is identified as nonjustified (1965). Extension of this logic to the variables of perceived violence, perceived reality or humor, or degree of enjoyment would be straightforward. Combining these factors would be a second stage in such a research sequence. We would project that among subjects for whom a given scene has been identified as low in violence, containing acceptable behavior, and which is considered humorous and real, maximum aggressive behavior should result.

The lack of replication of the factor of perceived violence in the present study cannot be overlooked. Among the ten-year-old blacks in the first study, there was a consistent perception of lesser violence. Indeed, it was one of the clearest and cleanest findings in that study. Among the 14-year-olds, there was no semblance of that pattern. We have no ready explanation for this finding. Because we used two age groups, rather than the same age group over four years, alternative explanations of generational *vs.* developmental differences remain inseparable. Although both groups were raised in a television environment, four additional years of watching television violence (combined with greater physical and psychological maturity) could contribute to the differences. Only subsequent research can identify the causal factors. Most immediately, parallel data among young girls should be acquired and compared with that found for the boys.

This research project adds further to the study of children and aggression. Perceptions of mediated violence are anticipated to be linked to the child's attitudes toward aggression and to his own aggressive behavior. Knowing responses to mediated violence among specifiable and major subgroups of the population, particularly among impressionable children, is expected to be critical in examining the general question of effects of television content.

Other implications of this research are discussed in our prior report. From these additional data, we can paraphrase the concluding comment of that paper: The evidence shows that the more disadvantaged are more aggressive in attitude and experience. To the extent that (a) this aggressiveness is strongly reinforced through a steady exposure to violent television fare, (b) such content is perceived as acceptable, liked, humorous, and less violent, and (c) few counteraggression messages are received from family, peers, or other socializing agencies, then the outcome has significant social implications.

FOOTNOTES

1. The research upon which this report is based was performed pursuant to Contract No. HSM 42-70-32 with the National Institute of

Mental Health, Health Services and Mental Health Administration, Department of Health, Education and Welfare. The research staff for this project included Tom Allen, Phil Ericson, Larry Kincaid, Mantha Vlahos, Esther White, and Edward Wotring. Professor Carol Smith, Western Michigan University, arranged for the use of public school students in Kalamazoo, Michigan. We are grateful to the public schools of that city for their cooperation in this and other projects.

REFERENCES

Berkowitz, L. Aggressive cues in aggressive behavior and hostility catharsis. *Psychological Review*, 1964, **71**, 104-22.

Berkowitz, L. Some aspects of observed aggression. *Journal of Personality and Social Psychology*, 1965, **2**, 359-69.

Chilman, C.S. Child rearing and family relationship patterns of the very poor. *Welfare in Review*, 1965, **3** (1), 9-19.

Clark, K.B. Dark ghetto: dilemmas of social power. New York: Harper & Row, 1965.

Gans, H.J. *The urban villagers: group and class in the life of Italian Americans.* Glencoe, Ill.: Free Press, 1962.

Greenberg, B.S., and Gordon, T.F. Social class and racial differences in children's perceptions of television violence. In *Television and social behavior*, Vol. 3 (preceding paper).

Hicks, D. Short-and long-term retention of affectively varied modeled behavior. *Psychonomic Science*, 1968, **11**, 369-70.

Himmelweit, H.T., Oppenheim, A.N., and Vince, P. *Television and the child.* New York: Oxford University Press, 1958.

McGuire, W. Attitudes and opinions. *Annual Review of Psychology*, 1966, **17**, 475-514.

Moles, O.C. Child training practices among low-income families—an empirical study. *Welfare in Review*, 1965, **3** (12), 1-19.

National Advisory Commission on Civil Disorders. *Report of the National Advisory Commission on Civil Disorders.* New York: Bantam, 1969.

Sears, R.R. Relation of early socialization experiences to aggression in middle childhood. *Journal of Abnormal and Social Psychology*, 1961, **63**, 466-92.

Siegel, S. *Non-parametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956.

Children's Violence Perception as a Function of Television Violence

Martin S. Rabinovitch, Malcolm S. McLean, Jr., James W. Markham, and Albert D. Talbott

University of Iowa

Investigation of the likelihood of a puzzling relationship between mass media content and instances of individual and collective violence has, for the most part, concentrated on the possible behavioral consequences of exposure to televised violence. In order to develop a more thorough understanding of this relationship, it seems useful to inquire into intervening perception processes—audience members' perceptions of media violence and their perceptions of the natural environment after exposure to media violence. The present study is devoted to the latter area of study, the influence of exposure to televised violence on children's everyday perceptions.[1]

Researching readiness to perceive violence is a relatively new way to study aggression. In a pioneer study, Toch and Schulte (1961) used a stereoscopic tachistoscope to present nine stereograms to three groups of students. Each stereogram contained a violent and a nonviolent figure. One figure was presented to the right eye, the other to the left eye. As a result of very short exposure, only one figure was seen. The dependent measure was the number of reports of violent content. Controls in the experiment (psychology students and first year police administration students) averaged 4.5 violent reports for 18 trials, while advanced police administration students averaged 9.4. Toch and Schulte suggested that this increased readiness to perceive violence was a result of training and experience which sensitized the advanced police students to violence.

Shelley and Toch (1962) next took the stereoscopic tachistoscope to a correction camp in Michigan and administered the perception task to young men between the ages of 17 and 24 serving sentences for theft and other offenses. High scores on the violence perception task formed the test group, which was matched for age and race with a control group. On the basis of an objective measure of institutionalized adjustment (being sent to a reformatory or to prison), the test group scored much higher than the control group. Subjects whose readiness to perceive violence was greater were more likely to end up in more severe institutions. In further examination, the case histories of the subjects were investigated. The experimenters found that high scorers on the violence perception task had histories of active violent behavior. Shelley and Toch predicted that perceptually violent individuals had a stable antisocial component which was manifested in sensitivity to violence and social acceptance of it. The low scorers on the violence perception task showed an inability to control themselves when influenced by others. They had histories of passive participation in delinquent activities.

In a more recent study, Moore (1966) added variables of age and sex. His findings were consistent with other aggression studies: males and older subjects were more ready to perceive violence.

The procedure developed by Toch and his associates was modified slightly to form one of the two measurements of violence perception used in the present study. Nine 35mm color slides of the violent behavior of two males and nine 35mm slides of the same males' nonviolent behavior replaced the silhouette-stereograms used by Toch. To permit multiple testing of observers, slides were projected on a screen using an alteration suggested by Toch (1970). Fitting the lenses of two slide projectors with polarizing lenses and giving the observers polarizing glasses, it is possible to expose the left eyes of a group of people to a violent slide while simultaneously exposing their right eyes to a nonviolent slide. Owing to the short exposure (one-half second), only the prepotent picture reaches awareness. Observers freely describe what they see, and the number of violent percepts is used as an index of readiness to see violence.

The stereoscopic procedure for measuring violence perception satisfies the four conditions set out by Tajfel (1969) to study social and cultural factors in perception: 1. They are directly tied to the sensory information received at the time of responding. 2. The response is not based on a chain of complex and abstract inferences. 3. The response does not involve acts of choice in which there is a clear awareness by the perceiver of alternatives to his perceptual decision. 4. The stimulus is of such a nature that it allows the possibility of one correct response concerning its relevant attribute (p. 321).

A second means of measuring violence perception was introduced to provide a more thorough investigation. Signal detection analysis can provide separate measures of an observer's ability to separate alternative stimulus classes and of the response bias he used in the decision process. With respect to violent and nonviolent stimulus classes, the perception process is divided into two subprocesses—ability to separate violent and nonviolent stimuli and readiness to report seeing violent stimuli. Ability to separate violent and nonviolent stimuli has one index, d', the index of signal detectability. Readiness to report seeing violent stimuli is stated in terms of values of the criteria or cutoffs the observer uses to decide whether or not to report the occurrence of a violent stimulus. [Green and Swets (1966) provide a more detailed discussion of signal detection theory, and Ulehla, Adams, and Dwyer (1969) provide a comprehensive application of signal detection theory to social stimuli. In the present inquiry, the signal detection analysis followed the model developed by Dorfman and Alf (1968, 1969). Dorfman, Keeve, and Saslow (in press) used this model in a study of ethnic perception. The procedure used in the present study followed that used by Dorfman et al. with the substitution of violent and nonviolent stimuli for the ethnic stimuli.]

At this time, speculative consideration of the relation between violence perception and violent behavior is appropriate. Studies mentioned earlier show that a violent background is a probable antecedent condition for proneness to see violence. Toch (1969) advanced the concept of "violence-proneness" based on experimental research and peer interviews with violent criminals. Toch's individual with violent habits and propensities is more ready to perceive violence than are other individuals: "Violence-prone people do not merely espouse violence as a doctrine or philosophy, but they tend to see the world in violent terms and respond to it accordingly" (p. 191).

The violence-prone conception is not a simple causal relationship, however. A dynamic interlocking relationship appears to exist between violence perception and violent behavior. Past violent experiences lead to a greater readiness to perceive violence, and, correspondingly, increased proneness to see violence leads to increased likelihood of actual violence. Theorizing from his interviews with violent criminals, Toch states:

> The probability of violence in violence-provoking incidents is tied to the extent to which the aggressor indulges in preclassification or selective perception. For instance, many of the participants in our police incidents scan human contacts assiduously for the possibility of threatening implications. They do so with varying degrees of intensity and with differing abilities to "spot" suggestive information. The actions of other people are eventually classified as either nonchallenging (safe) or as challenging (requiring action) (p. 185, 186).

The first stage of a violence-provoking incident—the perception of other people as menacing—appears to be influenced by the violence perception propensities of the aggressor. Other factors, like subsequent actions of other people and the behavioral habits of the aggressor, determine the eventual outcome of a given incident. However, proneness to perceive violence would seem to increase the probability of violent action in a violence-provoking incident.

Granting the influence of violence perception on violent action, we must ask whether perceptual biases obtained from natural experiences are equivalent to perceptual biases obtained from media experiences.

Siegel (1969) gives theoretical support for the equality of televised and actual violence. She feels that presentation of drama and news alternatively in the same medium contributes to a blurring, for audience members, of fact and fiction. Additionally, she says, television content is generally accepted as authentic owing to the vividness and fidelity of the medium: "Children understand such presentations as authentic and credible, and assume that the world is really the way it appears there" (p. 270). Siegel (1958) had pursued this line of thought in a role expectation study. She found that girls who had heard a certain radio episode, in which a taxi driver solved a problem by violent means, gave more violent descriptions when asked to complete a news story about local taxi drivers.

Televised violence is one of the many possible factors contributing to the development of children's proneness to see violence. One study, already described, has shown that past behavioral experiences with violence are related to proneness to see violence (Shelley and Toch, 1962). Additionally, Wolfgang and Ferracuti (1967) theorized that indoctrination into a violent subculture is an antecedent condition for proneness to perceive violence. Training experiences (Toch and Schulte, 1961) were shown to be related to violence perception.

A theory relating televised violence to violence perception is derived, in part, from general notions about perception. Perception is a process in which an individual's past and present experiences interact. An individual's past experiences give his present experiences their meaning.

Present experiences help interpret the future. Each experience contributes to a person's expectations about the future. Through his experiences, an individual forms assumptions about the world and acquires propensities to see certain elements of his environment. A participant in mass communication perceives mass media content with his past experiences, motivations, and predispositions all influencing this reception. Conversely, it is conceivable that mass media experiences can contribute to the perception predispositions of its consumers. The possibility arises, therefore, that violent television programs can foster a proneness in audience members to perceive violence.

The contribution of televised violence to violence perception is most likely governed by familiarity. Engel (1956) demonstrated that more familiar pictures of faces will predominate in binocular rivalry over less familiar pictures. Toch and MacLean (1962) presented the view that "those things that have been tied in most closely and most often with past personal experience predominate perceptually over the unusual or the unfamiliar" (p. 67). Toch and Schulte (1961) used the familiarity concept to explain the differences in violence perception between advanced police administration students and other students. Moore (1966) offered a similar line of thinking to explain males seeing more violence than females. Berkowitz (1962) stated that television can have a lasting influence on perception if the themes are repeated often enough.

In the present study, a group exposed to violent television content had relatively closer and more numerous experiences with violence than the other groups because of that exposure. When tested in an ambiguous situation, this group, which was more familiar with symbolized violence, would be expected to perceive more violence than groups not exposed to a violent television program. With respect specifically to the free-response, stereoscopic measure of violence perception, results in conjunction with the following reasoning put forward by Toch and Schulte (1961) would be expected:

> A momentary exposure of rival fields in a stereoscope presents a perceptual task in which one set of meanings must be elaborated at the expense of another. If the fields are mutually exclusive (so that they cannot 'fuse'), and if neither field asserts itself through structural advantages (such as those of a strongly articulated figure competing with a vaguely outlined one), familiarity clearly becomes the only remaining basis of choice (p. 392).

The following hypothesis was offered to test the above reasoning: When children previously exposed to a violent television program are presented with a series of simultaneous pairs of violent and nonviolent slides, they will see a greater number of violent slides than children previously exposed to a nonviolent television program or those exposed to no television program.

Parallel results were expected for the response-bias measure of the signal detection analysis. Observers' cutoffs, which indicate response biases, are influenced by the *a priori* probabilities and values and costs associated with the perceptual experience under consideration. Children who saw the violent television program should expect more violence and thus have a greater *a priori* probability for violent experiences. This should influence their response bias in the direction of a greater readiness to report seeing violent occurrences. The following hypothesis was offered to test this plausability: When children previously exposed to a violent television program are presented a randomly ordered series of 100 violent and 100 nonviolent slides, they will show more readiness to report a violent slide than children previously exposed to a nonviolent program or those exposed to no television program.

Since there was no content in the television programs used which would provide training in discriminating violence from nonviolence, no differences were expected to develop along the discriminability measure of the signal detection analysis.

Subsidiary hypotheses concerning differences in violence perception between boys and girls were established. No differences between boys' and girls' ability to discriminate violence were forecast. Boys do not appear to have training which would give them greater ability to discriminate between violence and nonviolence than females. But their differential socialization is more likely to be revealed as readiness to see violence and readiness to behave violently. Boys were expected to be more ready to see violence than girls owing to this differential socialization. Support for this view came from Moore's (1966) study. The following specific hypotheses were put forward: When presented with a series of simultaneous pairs of violent and nonviolent slides, boys will see more violent slides than girls. When exposed to a randomly ordered series of 100 violent and 100 nonviolent slides, boys will show more readiness to report a violent slide.

METHOD

Sixth-grade children of the Cedar Rapids (Iowa) School District who obtained parental permission participated in the study. Twenty-four girls and 33 boys took part.

Apparatus

Apparatus consisted of two videotape recorders with monitors, a Marietta projection stereoscopic tachistoscope, and polarizing "3-D" glasses. During the first half-hour of the experiment, two of three groups of children watched television programs on videotapes recorded during August 1970 from normal telecasts of two Cedar Rapids television stations. For this purpose, two videotape recorders with monitors (equivalent to the average portable television set) were set up in different rooms.

After the television viewing, violence perception was measured by using the tachistoscope and the polarizing viewers in the following ways:

1. The free-response procedure utilized both slide projectors of the stereoscopic projection tachistoscope. One slide projector projected a violent slide while the other was simultaneously presenting a nonviolent slide. The projector lenses were covered by polarizing filters set at 90 degrees to each other. Polaroid filters fitted in this manner caused the violent slide to be projected with light waves of different dimensions than the nonviolent slide. Observers (the children) wore polaroid glasses with lenses set at similar angles. During projection, two images, one violent and one nonviolent, were projected on a silver screen at the same time. (Use of a beaded screen would scramble the polarized light and thus destroy the desired effect.) Both violent and nonviolent images were concurrently present on the screen. If observers had not worn special glasses, each eye would have received two images. However, because of the polarizing filters, only one image reached each eye. A timer activated electronic shutters in front of each projector lens. Concurrent exposure of the projectors was set at .5 seconds throughout the free-response phase of violence perception measurement. At this short time interval, an individual could only perceive one of the two images.
2. Only one of the two slide projectors was used for the signal detection analysis of violence perception. Correspondingly, the polaroid lense was removed from this projector, and observers were asked to remove their polarizing glasses. Exposure was set at .02 seconds throughout 200 presentations involved in the signal detection procedure.

Mass media presentation

An episode of *Peter Gunn* was chosen as the violent program, because it was the only half-hour television program on the air at that time which had any substantial violent content. (*Peter Gunn* is no longer in production but is in syndicated distribution across the country.) An episode of *Green Acres* was chosen as the nonviolent program. Because it used actual television programs, the mass media presentation was thought to be representative of normal television fare. On the other hand, since the programs were presented just as they were broadcast, the type of violent content, the rewarding or punishing of violence, and other complications could not be controlled in this study. However, any individual who views the two programs would observe a lack of violent content in *Green Acres* and a fair amount of violence in *Peter Gunn*. Specifically, the *Peter Gunn* episode used showed: a convict hitting a guard over the head; a man holding a razor to Peter Gunn's neck; Peter Gunn involved in a fight, with gunfire, fist fighting, and glass breaking; one gangster shooting another, jumping through a glass window into the street, and finally being riddled with bullets by policemen. The *Green Acres* episode showed the hired hand placing a falsified advertisement in a lonely-hearts column, the exchange of letters between one respondent and the hired hand, and the humorous incidents which resulted from the surprise visit of the respondent.

Stimulus figures

Violent and nonviolent color slides were used in both measures of violence perception. A large number of slides (500) were made by photographing two males in various interpersonal activities. Violent activity was structured to conform to the following definition: "any situation in which fighting occurs or where one or both of the individuals attempts to hurt or injure the other or each other by any means." The violent set was comprised of slides meeting this definition. The nonviolent slides did not meet this definition.

Despite the fact that the slides were selected with the foregoing definition in mind, children viewing the slides for a relatively long period of time might not agree that a given slide did contain the content it was alleged to contain. To preclude any difference of opinion with respect to violent content by uninformed observers, validation of the stimulus figures had been carried out. Several months before the main experiment, 200 of the slides were shown to junior high school students in the eighth grade at the University of Iowa School. These Iowa City children were asked to score each slide using the preceding definition of violence and nonviolence. The validation consisted of showing each slide for a relatively long time, 30 seconds. Of the 200 slides, only two resulted in a substantial lack of agreement among the children. Both of these slides reputedly depicted verbal aggression. They were replaced by slides containing physical aggression. Replacement slides corresponded to other slides of physical aggression that had been validated. After this validation of stimulus materials, the techniques of measuring violence perception used in the present study conformed to Tajfel's (1969) fourth condition for studying social perception: "The stimulus is of such a nature that it allows the possibility of one correct response concerning its relevant attribute." Most individuals, children and adults, having the opportunity to view the slides for a relatively long time and armed with the established definition of violence, would agree that the slides called violent *were* violent and the slides called nonviolent *were* nonviolent.

The slides were employed differently in the two measures of perception. In the signal detection analysis, all 200 slides were used. In general, violent slides differed from nonviolent slides only in the form of behavior depicted. Surroundings, size of the pictures, material objects, and most other parameters of the violent slides were matched in the two slides. Only interpersonal behavior varied. For example, in one violent slide one man appears to be hitting the other over the head with an automobile jack. In the corresponding nonviolent slide, the same man is handing the other the jack.

In the free-response procedure, nine pairs of violent and nonviolent slides were chosen from the larger population of slides on the basis of the following criteria put forward by Shelley and Toch (1962):

1. One picture of each pair had to depict a crime-related or violent scene, whereas the other picture was selected as being devoid of violent content.
2. The two pictures had to be structurally matched so as to occupy the same space and location in the visual field.

Table 1: Description of pairs of slides used in free response measure of violence perception

Number	Violent slide	Nonviolent slide
1	One man hits the other over the head with a gun.	One man helps the other pound a pole in the ground with a gun butt.
2	One man shoots the other with a rifle.	Both men walk. One carries a rifle.
3	One man pushes the other off a bridge.	Both men walk on the bridge.
4	One man kicks the other off a merry-go-round.	Both men ride on a merry-go-round.
5	One man hits the other over the head with a book.	One man shows the other something in a book.
6	One man, tied up, tries to hit the other man.	One man holds a rope. The other looks at the rope.
7	One man hits the other over the head with a car jack.	One man helps the other take a car jack out of the car trunk.
8	One man hits the other over the head with a rock.	Both men help to lift a rock.
9	One man holds the other and takes money at gunpoint.	One man gives the other man money from his wallet.

3. The two pictures had to be mutually exclusive, in the sense of not fusing into a single image (p. 465).

Table 1 lists the content of the nine pairs of slides. Most slides used in the free-response procedure were also used in the signal detection method. However, in the free-response procedure, slides were presented in pairs consisting each of one violent and one nonviolent slide. In the signal detection method, slides were presented successively in a random order.

Procedure

Children whose parental permission had been obtained (73 percent of available children) were divided by sex and assigned to one of three groups by a table of random numbers. On the morning of the study, children went to their classrooms as usual. The elementary school where this study took place practiced an open-class approach to education. (This point is significant; as attested to by some teachers at the school, it is not uncommon for the sixth-grade class [equivalent in size to three traditional style classes] to divide into smaller section. Additionally, guests are not a rare occurrence at the school. Therefore, the procedure of dividing the students would not be likely to have created any special attention.)

Teachers of the sixth-grade class were given two lists of children. They sent one group to the library and another to the cafeteria. The third group remained engaged in normal school activities for the first half-hour with the children for whom parental permission had not been obtained.

When the children arrived in the library or cafeteria, they found the videotape recorder with a monitor equivalent to the average portable television set already set up. They were greeted and seated and then told, "We're here today to test out some new equipment. First we are going to watch a television program." The group in the library saw *Peter Gunn.* The group in the cafeteria saw *Green Acres.* Just before the television programs ended, a list containing the names of the children in the third group was sent to the open classroom. An assistant accompanied this third group to the cafeteria. At this time there were *four* groups of children: (1) girls and boys who saw *Peter Gunn,* (2) girls and boys who saw *Green Acres,* (3) girls and boys who engaged in normal school activities, and (4) those children, girls and boys, for whom we had no parental permission, who did not participate in the experiment. All reference to "groups of children" will be understood to refer only to the first three groups.

After the television presentations, children were given coded response sheets, given the opportunity to use restroom facilities, and asked not to communicate verbally with any of their fellow students.

Children who were already in the cafeteria were spread throughout the room. As the two other groups of children arrived, they were evenly distributed to ensure no bias in location for any particular sex or group. Once all the children were seated, each child was given a pair of polarizing glasses. The children were told that they were to help see how well a new type of slide projector works at fast speeds. They were encouraged to tell the truth and told that no one at the school would know what they as individuals had written down. During the next half-hour, the time was devoted to the free-response, stereoscopic method of measuring violence perception. The children saw each of nine pairs of slides and were asked to write a short description of each presentation. The nine pairs were then repeated in an alternate order to rule out the effects of eye dominance. After the eighteen presentations, each slide had been exposed to each eye once.

Observers removed their polarizing glasses and rested for a short time. They were then instructed how to respond to the next set of slides. They were told that after each slide they were to ask themselves two questions and answer by making two "dashes" on the response sheet. The verbatim instructions were as follows:

> First ask yourself, "What happened?" Each slide shows two men doing something. If there is fighting, or if one or both men tries to hurt the other in any way by using any part of the body, including yelling or by using any object like a stick or a gun, place a dash under "hurting or fighting!" Then ask yourself, "How sure are you?" If you are pretty sure, place a dash under "very sure." If you are not sure or are just guessing, place a dash under "not very sure."

Analysis of the free responses to the slide pairs consisted of scoring the written descriptions observers made of each slide. A description was given one point if it was clearly violent, for example, "Two men fighting." A description was given a half point if the violent slide was described but not in positively violent terms. For example, in response to the slide pair of (1) two men picking up a rock, and (2) one man hitting the other over the head with a rock, an observer wrote, "2 men, 1 boy throwing leaves at 1 man." A description was given no points if the nonviolent slide was described or if the description was nonviolent. Very few descriptions were ambiguous enough to receive a half point. 99.32 percent of the picture descriptions received either one point or no points.

Signal detection analysis was accomplished by transferring each observer's responses from the response sheets to computer cards. A computer program was then used to compare each observer's response on each slide with the correct response and then to compute the proportion of responses in each cell of a response matrix.

Proportions were tabulated in eight cells. The first four cells were for responses to nonviolent slides. If a slide was nonviolent, the response alternatives were: (1) "no hurting or fighting, very sure"; (2) "no hurting or fighting, not very sure"; (3) "hurting or fighting, not very sure"; and (4) "hurting or fighting, very sure." If a slide was violent, the response alternatives were: (5) "no hurting or fighting, very sure"; (6) "no hurting or fighting, not very sure"; (7) "hurting or fighting, not very sure"; and (8) "hurting or fighting, very sure." Cells one, two, seven, and eight would be correct responses or hits. Cells three, four, five, and six would be incorrect responses or misses.

RESULTS

Table 2 contains a summary of the mean violence perception scores. An analysis of variance of the free-response measure of violence perception revealed that the difference between the groups who viewed different television programs was significant ($F = 7.704$, $df = 2/51$, $p < .01$). Neither the difference between sexes nor the interaction between sex and television presentation was significant. Followup contrasts revealed that children in the violent television exposure condition were least ready to perceive violence when measured with the free-response measurement. Children in the nonviolent television condition were centrally located in terms of readiness to perceive violence. Children who did not see any television program were most ready to perceive violence for the free-response measure.

Table 2: Mean number of violent pictures seen by children in free response measure of violence perception

Sex	Television exposure		
	Violent program	Nonviolent program	No program
Males	1.25	2.41	2.40
Females	.81	2.19	3.38
Mean over both males and females	1.08	2.31	2.83

Signal detection analysis of the children's responses to the 200 presentations of violent or nonviolent slides resulted in more detailed information about the children's perception of violence. This information is tabulated in Table 3. For the discriminability subprocess, a computer program developed by Dorfman and Alf (1969) was used to obtain indices of discriminability for each observer. An analysis of variance of the discriminability indices did not reveal any significant differences between television conditions, between sexes, or among interaction of conditions and sexes. For the decision subprocess, the same computer program was used to obtain indices of the children's cutoffs. No significant differences resulted with respect to cutoff one.

Table 3: Mean scores of children on signal detection measures of violence perception

	Perception Measure				
	Discriminability d'	Decision bias (readiness to report violence)			
Exposure		Cutoff 1[a]	Cutoff 2[b]	Cutoff 3[c]	Overall probability of reporting violence
			Males		
Violent program	0.982	–0.535	0.290	0.760	0.550
Nonviolent program	1.530	0.014	0.756	1.157	0.476
No program	1.100	–0.246	0.427	1.016	0.537
			Females		
Violent program	1.317	–0.423	0.427	1.003	0.522
Nonviolent program	1.170	–0.567	0.562	1.354	0.486
No program	1.194	–0.247	0.630	1.290	0.475
		Mean over both males and females			
Violent program	1.116	–0.494	0.340	0.850	0.538
Nonviolent program	1.370	–0.252	0.670	1.244	0.480
No program	1.142	–0.246	0.517	1.137	0.509

[a] Larger, more positive value indicates greater readiness to report "no hurting or fighting , very sure".

[b] Larger, more positive value indicates less readiness to report "hurting or fighting" as opposed to "no hurting or fighting".

[c] Larger, more positive value indicates less readiness to report "hurting or fighting, very sure".

Cutoff two indicated the tendency to report "no hurting or fighting, not very sure" as opposed to "hurting or fighting, not very sure." With respect to cutoff two, results opposite to those of the free-response measure appeared. The variability owing to television exposure was significant (F = 5.092, df = 2/49, p < .05). The order of groups in terms of proneness to report violence on cutoff two was (from greatest to least): violent program, no program and nonviolent program. The order of television conditions with reference to cutoff three was exactly the same as that of cutoff two and was significant (F = 5.410, df = 2/49, p < .01). In addition, cutoff three was the only measure which revealed significant differences between males and females. As expected, males were more ready to report "hurting or fighting, very sure" than to report "hurting or fighting, not very sure" (F = 6.158, df = 1/49, p < .005).

Overall probability of reporting violence, an auxiliary index of response bias, failed to bring out any significant differences.

DISCUSSION

The use of two methods of violence perception measurement resulted in a nonuniform pattern of results which suggests the operation of a number of variables. Of the possible factors which could have influenced violence perception, familiarity was expected to play the largest role. However, results opposite to this hypothesis emerged. Groups who were less familiar with violence displayed more readiness to see violence as indexed by the free-response measurement of violence perception. To clarify the results obtained with the free-response measurement, alternate explanation is necessary.

One explanation, relating to the values and costs associated with perceiving violence, is perceptual defense. According to this concept, a response bias exists against reporting the perception of a socially disapproved action like violence. Toch and Schulte (1961) entertained perceptual defense as an alternative explanation for their experimental results. They stated that it was possible that law enforcement training removed inhibitions against perception of antisocial behavior. Perceptual defense is the perceptual counterpart of the concept of inhibition, which has been employed in many studies of overt aggression. In terms of values and costs, the costs associated with the perception of a socially disapproved action are greater than those associated with perception of a socially approved action. But the differences between the no-program and nonviolent program groups does not follow from this perceptual defense explanation.

Another possible explanation is derived from the frustration-aggression hypothesis. According to this interpretation, television programs can lessen frustration by closing the gap between expected and actual need satisfaction. A number of studies dealing with political and social violence (Gurr, 1967, 1968, 1969; Davies, 1962, 1969; Feierabend and Feierabend, 1968, 1969) have shown that frustration—a gap between wants and satisfactions—is the prime condition for collective violence. Frustration, in this sense, can be modified either by changing expectations or by changing satisfactions. Feierabend and Feierabend (1968, 1969) used the number of receivers of certain mass media per unit of population as part of their index of satisfaction.

Television programs are thought capable of providing satisfaction for audience members. One possible conception of this satisfaction follows Berlyne's (1960, 1967) theory that an individual will strive to satisfy a desire for an optimal amount of arousal potential in incoming stimuli. Depiction of violent content in the mass media is presumed to be a means for satisfying this desire for arousal. Of course, a number of factors determine the amount of arousal potential in a television program. Berlyne feels that the arousal potential of incoming stimuli is a function of intensive variables, affective variables, and collative variables (change and complexity, conflict and uncertainty). A nonviolent program which has an optimal degree of change, complexity, and uncertainty will most likely be more pleasing than a violent program which has little of these qualities. But, all other variables being equal, violent content is assumed to have an advantage over nonviolent content in satisfying audience members. One possible reason for this advantage is that violence is an extreme form of conflict, and ". . .complexity or diversity excites through some form of conflict" (p. 233).

This explanation further assumes that an individual who has a relatively large gap between expected and actual satisfaction of this desire for optimal arousal will be more frustrated and consequently be more prone to see violence than an individual who has less of a gap between expected and actual satisfaction of this desire. Children exposed to *Peter Gunn* were possibly less frustrated than the other groups. They are thought to have had the least gap between expected and actual need satisfaction. Exposure to *Green Acres* was potentially more frustrating than exposure to *Peter Gunn*, but less frustrating than exposure to no television program. Therefore, immediately following exposure to the television programs, the order of groups in terms of greatest readiness to see violence should have been (and was): no program, nonviolent program, and violent program.

Catharsis is the final supplementary concept we will examine. The catharsis concept differs from the frustration-aggression formulation in that proponents of catharsis emphasize satisfaction of *aggressive impulses* rather than the satisfaction of curiosity or arousal desires. Feshbach (1969) maintains that ". . .vicarious aggressive experiences serve to help regulate and partially satisfy aggressive tendencies" (p. 470). Reasoning directly from the catharsis hypothesis, the differences between the no-program and nonviolent program groups would not have been anticipated. The frustration-aggression formulation would anticipate this difference and, for this reason, appears to be the most useful explanation of the free-response results.

(As we have already stated, the signal detection results ran counter to the free-response results. Not much weight can be placed on the signal detection results, because the design of the study was such that the effect of time and the effect of different perception measurement cannot be separated. Further research is needed to unravel time and perception measurement. Such a study is now in progress.)

If the influence of time between the first measure and the second measure was nonexistent, the conflicting results can be explained in terms of perceptual defense. The first measure of violence perception did not appear to be as socially permissive of violent responses as the second measure. In the first measure, the free-response measure, children were asked to describe what they saw. Children were asked in the presence of an adult during school hours. Levin and Turgeon (1957) have shown that a strange adult can have an inhibiting influence on children's aggressive behavior. In the second measure, children were given the two categories "hurting or fighting" and "no hurting or fighting." Past studies (Siegel, 1957; Siegel and Kohn, 1959) have shown that an adult can decrease aggression anxiety by defining the social situation in favor of the propriety of aggression. Perhaps the social situation during the free-response measure tended to provoke more anxiety concerning aggression than the situation during the signal detection measurement.

Considering this observation, perceptual defense can partially illuminate the results. According to this explanation, children who saw *Peter Gunn* had more aggressive tendencies aroused than the other groups. However, the corresponding anxiety associated with violent percepts was subsequently higher for the violent group during the first measure of violence perception. During the second measure of violence perception, the social situation included approval of seeing violent percepts and perceptual defenses disappeared. This proposal cannot, however, account for differences found between the nonviolent program and no-program groups and requires additional study of the influence of time.

If the perception measures were equivalent, the conflicting results can be explained by combining frustration and familiarity. Possibly frustration was the main influence for approximately one hour after exposure to the television programs. However, as time went on, satisfaction obtained from watching television began to dissipate while the influence of familiarity, which was previously overshadowed by frustration, became prominent. This explanation requires supporting evidence to determine the equivalence of the perception measures.

SUMMARY AND CONCLUSION

The probability of violence in a given situation is tied to the extent to which the individuals involved are ready to perceive violence. To see how television violence affects this readiness to perceive violence, the investigator exposed sixth-grade children either to a violent or nonviolent television program or to no television program. After the children participated in one of these conditions,they evaluated violent and nonviolent slides flashed on a screen. Two measures of violence perception were used. In one measure, adapted from the work of Toch and his associates, violent and nonviolent slides were simultaneously flashed on a screen by means of a stereoscopic projection tachistoscope. Each eye of each observer received a separate image, one violent, the other nonviolent. Owing to the short time of exposure, only one image reached awareness each time. Observers freely described what they saw, and their propensity to perceive violence was determined by the number of violent percepts. In a second measure, signal detection theory was applied to the children's responses. Slides were flashed, one at a time, on the screen by one projector of the tachistoscope and observers were instructed to make two responses on each trial: "check either 'hurting or fighting' or 'no hurting or fighting' and check either 'very sure' or 'not very sure'." These responses were analyzed and yielded two measures of violence perception: readiness to report perceiving violence (similar to the first measure) and ability to discriminate violent from nonviolent experiences.

Did the children who saw the television program *Peter Gunn* then see the world in more violent terms than the other groups? The answer depended on whether violence perception was measured with the stereoscopic, free-response technique or with the signal detection analysis. With the free-response measure, children exposed to *Peter Gunn* were least ready to see violence. Children exposed to a nonviolent program, *Green Acres*, were more ready to see violence than children who saw the violent program, but less ready to see violence than the group who was not exposed to a television program. In a subsequent measurement, children exposed to a violent program changed from least to most ready to perceive violence. In both measures of perception, children who were not exposed to any program showed more readiness to perceive the world in violent terms than did children exposed to a nonviolent program. A minor finding uncovered, but not consistently across both measures, was that males were more ready to classify events as definitely violent than were females.

To explain the results from the free-response, stereoscopic measure of violence perception, a number of concepts were set forth. The most useful explanation employed frustration of a desire in audience members for optimal arousal. Children exposed to *Peter Gunn* were presumed to be least ready to see violence during the free-response measurement because they were least frustrated. They were assumed to have the smallest gap between expected and actual satisfaction of a desire for optimal arousal. Their exposure to *Peter Gunn* provided the most satisfaction of this desire. Accordingly, the *Green Acres* group was more frustrated (less satisfied) than the *Peter Gunn* group, but less frustrated than the no-television group. The preceding possible influence of violent programs must be tempered by the conflicting results of the signal detection analysis.

One implication suggested by the results of this study is that exposure to television content, both violent and nonviolent, can produce changes in children's perception of violence. The extent of this generalization requires moderation since no provision was taken to study the relative influence of other factors besides television. Berkowitz (1962) feels that mass media content is not as powerful a pressure on youngsters' perception as their parents and friends are. Parents potentially would exert more influence in a normal television viewing situation. Children in the present study viewed a television program in groups of peers rather than in a family situation. Even so, a lasting influence of television content on perception could result from repeated exposure to the same type of television programs.

About the nature of the influence of television content on violence perception, even more modest inferences must be made. The only consistent finding was that children in the created situation who were exposed to a particular nonviolent television program were less ready to see violence than children not exposed to any television program. Exposure to programs similar to *Green Acres*, under the limited set of circumstances used here, likely has the capability of influencing violence perception in the direction of less readiness to see violence. However, repetition of exposure and pressures of parents and peers are two variables already mentioned which limit this generalization. In addition, qualification is in order because of the noninclusion of several other variables. In designing the present study, it was decided to place the emphasis on measuring violence perception. Two measures of violence perception were used instead of including auxiliary variables.

One of the most interesting dimensions not covered is the possible differential influence of televised violence on girls and boys of different backgrounds and different behavior propensities. More specifically, dissimilar results might have been found for children who were high aggressive as opposed to low aggressive, for children who were delinquent as opposed to nondelinquent, and for children who came from middle-income as opposed to lower-income backgrounds. Through averaging, differential pressures on high and low aggressive children might have been hidden. On the other hand, since males are generally considered to be more aggressive than females, the lack of any interaction here between the variables of sex and television exposure suggests that violence perception of high aggressive children will not be influenced differently than that of low aggressive children by exposure to violent or nonviolent television programs. Even so, children's levels of general aggressiveness are likely to be an important factor in the contribution of televised violence to subsequent violence perception. Meyerson (1966) found that low aggressive children were affected more by a specific violent film than were high aggressive children when tested for violent behavior. The inclusion of level of general aggressiveness, then, is a most promising augmentation of media violence perception research.

Generalization of results is further limited to middle-class children in a relatively progressive school in the midwest United States.

The most fragile implications of the present study deal with the influence of violent television programs, in general, on violence perception of children. The results leave an inference that exposure to a violent television program decreases readiness to see violence for about an hour after exposure. It is possible that this influence changes during the second hour after exposure. However, before any conclusion can be reached concerning the effect of time on violence perception, additional research must be undertaken to separate the type of perception measurement from the time of measurement. A number of limitations of the preceding inference from the free-response measurement have already been mentioned. The signal detection analysis revealed contrary results. The exclusion of manipulation of level of general aggressiveness, of measurement over a wide range of time, and of the influence of parents and peers restricts this generalization about violent content.

The explanation given for the results requires corrobarative testing. Specifically, inclusion of an index of satisfaction and frustration is a needed extension of the present inquiry.

The implied contribution of groups exposed to violent, nonviolent, and no television is limited because only one sample of each exposure condition was used. Finer differentiation of exposure conditions might have uncovered dissimilar operation of various types of nonviolent program, violent program, and no-program exposures. Distinctive results would be anticipated for additional variables such as justified versus nonjustified violent content, rewarded versus punished violence, and content with surroundings similar to those used in the perception test versus dissimilar surroundings. A wide variety of nonviolent program and no-program exposure conditions could also lead to diverse results. In sum, a broader range of program samples is a logical extension of this inquiry.

This study has only begun to explore the relationship between exposure to violent television content and subsequent violence perception. Additional investigation is needed before a complete picture emerges of the complex interrelationships among mass media content and perceptions of its audience. Some of the more likely prospects for additional research have already been mentioned. Most needed is a study separating the effects of time and of different perception measurements. In conjunction with this recommended investigation, work should be initiated which couples perception of television programs during viewing with later perception in the nontelevision environment and which connects both of these factors with behavioral consequences of violent television content. While these qualifying factors should be kept in mind, the present study did demonstrate that televised violence can influence violence perception. The exact direction of this influence was nonuniform in the present study.

FOOTNOTES

1. This research was supported by Grant No. HSM-42-70-49 from the National Institute of Mental Health, United States Department of Health, Education, and Welfare. The author wishes to thank his wife Marilynne for her assistance throughout the study. The author is greatly indebted to Dr. Donald Dorfman for his help in designing the signal detection methodology and analysis and for the use of his computer program. Dr. Leonard D. Eron is thanked for his advice. Dr. George Ross, director of Research for the Cedar Rapids Unified School District, is thanked for his assistance in coordinating the study in Cedar Rapids. Thanks are extended to Mr. Larry Wood for constructing a computer program to calculate probabilities from individual ratings for the Dorfman and Alf computer program. Additional thanks are extended to Dr. Steve Haisman and Mr. Harley Straus for their voluntary assistance in photographing the slides, to Dr. George Gunkle for his careful critique of the manuscript, and to the following who contributed their assistance or suggestions to the project: Mr. Joe Carder, Miss Marian Elmquist, Mr. John Cortiss, and the administrators, faculty, students, and parents of the schools that helped with this study.

   Dr. Rabinovitch is now at San Fernando Valley State College, Department of Radio-TV-Film.

REFERENCES

Berkowitz, L. *Aggression: A social psychological analysis.* New York: McGraw-Hill, 1962.

Berlyne, D. E. *Conflict, arousal, and curiosity.* New York: McGraw-Hill, 1960.

Berlyne, D. E. Arousal and reinforcement. In D. Levine (Ed.) *Nebraska Symposium on Motivation.* Lincoln: University of Nebraska Press, pp. 1-110.

Davies, J. C. Toward a theory of revolution. *American Sociological Review,* 1962, **27**, 5-19.

Davies, J. C. The J-Curve of rising and declining satisfactions as a cause of some great revolutions and a contained rebellion. In H. D. Graham & T. R. Gurr (Eds.) *The history of violence in America,* New York: Bantam, 1969, pp. 670-730.

Dorfman, D. D., and Alf, E. F., Jr. Maximum likelihood estimation of parameters of signal detection theory—a direct solution. *Psychometricha.* 1968, **33**, 117-124.

Dorfman, D. D., and Alf, E. F., Jr. Maximum likelihood estimation of parameters of signal detection theory—rating method data. *Journal of Mathematical Psychology,* 1969, **6**, 487-497.

Dorfman, D. D., Keeve, S., and Saslow, C. Ethnic identification—a signal detection analysis. Submitted to *Journal of Personality and Social Psychology*, 1970.

Engel, E. The role of content in binocular resolution. *American Journal of Psychology*, 1956, **69**, 87-91.

Feierabend, I. K., and Feierabend, R. L. Conflict crisis and collision: A study of international stability. *Psychology Today*, May 1968, 26-70.

Feierabend, I. K., Feierabend, R. L., and Nesvold, B. A. Social change and political violence: cross-national patterns. In H. D. Graham and T. R. Gurr (Eds.) *The history of violence in America*, New York: Bantam, 1969, pp. 632-687.

Feshbach, S. The catharsis effect: Research and another view. Appendix III-E in R. Baker, and S. Ball (Eds.) *Mass media and violence: A report to the National Commission on the Causes and Prevention of Violence*. Washington: U. S. Government Printing Office, 1969.

Green, D. M., & Swets, J. A. *Signal detection theory and psycho physics*. New York: John Wiley, 1966.

Gurr, T. R. Explanatory models for civil strife using aggregate data. Paper presented at the annual meeting of the American Political Science Association, Chicago, September 1967.

Gurr, T. R. Psychological factors in civil violence. *World Politics*, 1968, **20**, 245-78.

Levin, H., and Turgeon, V.F. The influence of mother's presence on children's doll-play aggression. *Journal of Abnormal and Social Psychology*, 1957, **55**, 304-08.

Meyerson, L. J. The effects of filmed aggression on the aggressive responses of high and low aggression subjects. Unpublished doctoral dissertation, University of Iowa, 1966.

Moore, M. Aggression themes in a binocular rivalry situation. *Journal of Personality and Social Psychology*, 1966, **3**, 685-88.

Shelley, E. L. V., and Toch, H. H. The perception of violence as an indicator of adjustment in institutionalized offenders. *Journal of Criminal Law, Criminology, and Police Science*, 1962, **53**, 463-69.

Siegel, A. E. Aggressive behavior of young children in the absence of an adult, *Child Development*, 1957, **28**, 371-78.

Siegel, A. E. The influence of violence in the mass media upon children's role expectations. *Child Development*, 1958, **29**, 35-36.

Siegel, A. E. The effects of mass violence on social learning. Chapter 12 in R. Baker and S. Ball (Eds.) *Mass media and violence: A report to the national commission on the causes and prevention of violence*. Washington: U. S. Government Printing Office, 1969.

Siegel, A. E., and Kohn, L. G. Permissiveness, permission, and aggression: The effect of adult presence or absence on aggression in children's play. *Child Development*, 1959, **30**, 131-41.

Tajfel, H. Social and cultural factors in perception. In G. Lindzey and E. Aronson (Eds.) *The handbook of social psychology* (Vol. 3, *The individual in social context).* Reading, Mass. Addison-Wesley, 1969.

Toch, H. H. *Violent men: An inquiry into the psychology of violence.* Chicago: Aldine, 1969.

Toch, H. H. Personal communication, 1970.

Toch, H. H., and MacLean, M. S., Jr. Perception, communication and educational research: A transactional view. *AV Communication Review*, 1962, **10**, 55-77.

Toch, H. H., and Schulte, W. H. Readiness to perceive violence as a result of police training. *British Journal of Psychology*, 1961, **52**, 389-93.

Ulehla, Z. J., Adams, D. L., and Dwyer, J. G. Detection theory, Psychologically defined stimulus alternatives, and social learning theory. *Proceedings of the Annual Meeting of the American Psychological Association*, 1969.

Wolfgang, M. E., and Ferracuti, F. *The subculture of violence; towards an integrated theory in criminology.* London: Tavistock Publications, 1967.

Violence and Color Television: What Children of Different Ages Learn

Natan I. Katzman

Michigan State University

This research project was conceived in the midst of the controversy over the effects of mass media—especially television—on children. The basic question can be posed simply: Does aggressive violence on television increase aggressively violent behavior? Unfortunately, a complex series of factors influence the effects of mass media, and no simple answer can be given. Children of different sex, age, race, or background may respond differently to the same programs. The environment before, during, and after exposure may produce different types of responses. The context within the program where violence is used may also affect viewer behavior. With all these factors in mind, we have looked at learning—at what children recall from a program.[1] Perhaps the effects of certain variables on learning from television can provide clues to the effects of these variables on subsequent behavior.

This was a limited project. The subjects studied were drawn from a narrow demographic category. The stimuli used were two versions of a single program. Thus, we should be cautious about generalizing from the findings. However, since the topic is important, even the findings of a limited project are better than no data at all. They provide a starting point for further theory and research.

Violence, the use of color format, and the age of the viewer were the three main variables studied in this project. If high violence on television has an effect on behavior, then effects on recall might also be expected. Without entering the debate between the social learning theorists (who argue that aggressively violent media content is learned and imitated) and the exponants of the catharsis hypothesis (who claim that viewing aggressive violence reduces the tendency to behave that way), we can simply note that while a great deal of research has looked at behavioral effects of violent content, not much has been done on recall of program content. Color television is already a fact of everyday life in about half of all homes in the United States. Since there is concern with the effects of violence, it seems important to determine if these effects are the same when content is shown in black and white or in color. One of the variables receiving some attention at this time is the effect of "consequences" (does the viewer see the effect of aggressive violence on the victim). If the consequences of violence have an effect on viewers, there is a distinct possibility that the use of color to heighten "realism" compounds the effect. There is also the possibility that color television attracts more attention to violence and thus enhances effects. The third main variable in this study was the age of the young viewer. If different age groups respond differently to violence on television, a pause might be in order before conclusions are drawn from studies that have looked at only one group.

WHAT DOES THE PRIOR RESEARCH SAY?

A search of the literature produced no useful prior research regarding the effects of violence on recall of program content. However, some studies provide insights into the effects of color on learning and the differences in learning for children of different ages.

Incidental learning by children from the mass media

In recent years there has been growing evidence to support the proposition that children develop the ability to learn relevant material and incidental material in two independent patterns. For example, Roberts (1968) found that younger children recalled more about color slides of a speaker, relative to what they recalled about his recorded message, than older children. Similarly, Collins (1970) and Hale, Miller, and Stevenson (1968) report that as a child grows older he can recall more central material, while there is a curvilinear relationship between age and recall of peripheral material. These findings have been interpreted to mean that "focusing" increases with age. The explanation is that very young children attend equally to both relevant and irrelevant aspects of a message, as they get older their ability to recall both types of material increases, and as they get even older (above 12 years) they develop the ability to ignore peripheral material by focusing on the central aspects of the message.

The data tend to indicate that there are two different learning modes and that these modes develop differently. In younger children, learning of detail—both relevant and irrelevant—seems to be a more developed skill than ability to conceptualize a total message while discounting extraneous factors. On the other hand, the former ability seems to improve less rapidly than the latter. Thus, older children, while able to learn detail somewhat better than they could have a few years before, are *more* improved in their ability to attend to relevant content and discount irrelevant material.

Issues such as these become socially important in the context of other research which has been conducted or is currently underway. Evidence pertaining to the effects of such variables as perceived justification, reward, social approval, etc., is being collected in an attempt to determine the effects of mass media on behavior, especially violent behavior among children. (The work for the Television and Social Behavior program (1971) of Chaffee at the University of Wisconsin, Gerbner at the University of Pennsylvania, Roberts and Leifer at Stanford University, and Greenberg at Michigan State all will provide evidence on this matter.) Variables of this type are only operative to the degree that a stimulus program containing any of them is perceived by the viewer to be a coherent whole. Higher levels of learning of central material, then, would facilitate the operation of such factors. Attention to details (as opposed to total content) would raise the functional importance of isolated, specific occurrences within a program. For example, in a program where a character is seen to resort to violence in order to obtain his ends (a common occurrence, according to Larson, Grey, and Fortis, 1963), the consequences of this violence and his punishment, if any, would be more relevant to viewers with a tendency to view the episode as a whole. On the other hand, the violent act—relatively free of qualifying factors—would be more salient to viewers who learned more about detail than overview. Hence, the study of incidental learning can provide insight into aspects of television-mediated behavior.

Prior research indicates that younger children may pay less attention to the overall program content. They would be most affected by the isolated instances of aggressive violence regardless of context. Older children may pay more attention to the central plot and theme of violent programs. They may be more affected by such variables as justification, reward, consequences, and motivation.

The effects of color television

At present the use of color television in the United States is growing at an accelerating rate. If the pattern of growth follows the typical S-curve for diffusion of new technologies, we can expect color sets in over 90 percent of the homes in this country within the decade. This phenomenon raises the importance of evidence concerning differential effects of color, for current studies of television effects have almost exclusively studied black and white formats. There is the possibility (which may not be as remote as it might first seem) that studying the effects of violence on black and white television will have about as much social significance as studying the effects of violence in radio drama. Soon, *color* television will be the single most significant mass medium, and it may not have the same effects as black and white television.

Prior research has indicated that variations in modes of communication, including the use of color, produce effects that are independent of the effects of content. Deutschman, Barrow, and McMillan (1961) found some evidence that a live presentation produced a higher ratio of "irrelevant" to "relevant" learning than a film presentation of the same material. Extraneous visual cues (although not necessarily color cues) in the live presentation were said to cause an increase in attention to what was called "irrelevant" material. A study by Kumata (1960) has a more direct bearing on the effect of color presentations. He compared two forms of an advertisement: one projected in color on a large screen, the other shown in black and white on television. "Details" were better retained by subjects who saw the color/projected version, and "principles" were better retained by subjects who saw the televised black and white version. Schaps and Guest (1968) reported similar findings in another study of advertisements. They inserted color and black and white versions of the same commercials into filmed versions of a television show. More "details" were remembered from the color version.

Two of the most rigorous studies of the effects of color, by Vandermeer (1954) and Kanner and Rosenstein (1960), indicated that color format had no appreciable effect on recall of nonverbal (visual) or verbal material in a training situation. Although they did not specify such a distinction, both of these studies seem to have tested retention only for central material and made no attempt to judge the effects of color on learning of peripheral material. Yet the findings are strong and consistent. The addition of the cues provided by color format have no clear effect on recall of central material.

A recent study by Katzman and Nyenhuis (1971) has indicated similar effects of color on learning from an entertainment medium. They presented a story from a comic book photographed on 35 mm slides. Subjects who saw the story in color recalled significantly more nonverbal peripheral material than subjects who saw the story in black and white. The color and black and white presentations produced equal amounts of learning for material that was verbal or central to the story. In sum, it seems that over various media the effect of color presentation is an increase in the ratio of peripheral to central material recalled from the visual channel.

DEFINITIONS, REASONING, AND HYPOTHESES

The central peripheral distinction

It should be clear by now that a major focus of this study was the differences in scores on recall of central and peripheral material. Various headings have been given to this type of distinction: "incidental learning" and "learning of details" are both phrases that seem to be closely related to the notion of peripheral, or irrelevant, material. Unfortunately, the notion of incidental learning has been used in two distinct ways. Some experimenters consider "incidental" material much the same as we shall define peripheral content; but a major strand of research has defined incidental learning as learning that takes place when a person has no incentive to learn, learning from a message that was not intended to inform. The latter type of definition is not exactly within the focus of interest in this project. Similarly, the notion of "details" can be used in operations that seem to test the central/peripheral distinction; but details can be either central or peripheral by the definitions we shall use. Frequently the operationalization of the concept (e.g., Schaps and Guest, 1968) is actually a test for *irrelevant* details. However, the notion of "details" seems too ambiguous for our purposes.

Collins (1970) defined *central material* as *"essential to the narrative sense of the presentation."* This definition seems to hold the key to the distinction between central and peripheral material. *Central material is program content that is relevant to the basic information, message, plot, or theme.* Content that is central is important. It is a key part of the total presentation and gives information about the main events in a program and their relationships.

Peripheral material is the complement or residual of central material. *Program content that is not relevant to the basic information, message, plot, or theme is peripheral material.* Anything that can be identified within a presentation does not meet the criteria for central material can be considered peripheral material.

The utility of the central/peripheral distinction has already been demonstrated in the study of the cognitive development of children and in research on the effects of color format on learning. It seems that such a distinction may also prove useful in the study of the effects of media content.

Aggressive violence

In another paper (Katzman, 1971), I have attempted to discuss distinctions between aggression and violence. It seems that the major concern over possible mass media effects is, and has been, the fear that aggressive violence in the media produces aggressively violent behavior. *Aggressive violence is behavior intended to physically harm another party.* The viewer of aggressive violence on television must be aware that one person intends to physically harm another person. If these criteria—perceived intent, potential physical harm, human victim—are present, an incident on television can be called aggressively violent.

Two hypotheses were made concerning the effects of aggressive violent content on learning from a program designed to entertain:

H_1: *More total material will be learned from a program that is relatively high in aggressively violent content.*

H_2: *The difference between peripheral material learned and central material learned will favor peripheral learning if a program is high in aggressive violence.*

The reasoning behind these hypotheses is derived from the notion that aggressive violence is used as a device to draw viewer attention to a program. (Baldwin and Lewis, 1971, have reported that the men who make television "action" programs feel that the use of violence increases the appeal of programs.) Increased attention might be expected to result in more overall recall for program material. It might also be primarily directed toward peripheral material in the programs. Thus, while total learning would increase, the learning of peripheral material would increase relative to the learning of central material.

Color presentation

The main hypothesis concerning the effect of color presentation was based on the earlier research on color effects in media other than television. Color had not been found to affect the recall of central material, but it did seem to affect the recall of peripheral material in the visual channel.

H_3: *The difference between peripheral-visual material learned and central-visual material learned will favor peripheral-visual material if the presentation is in color.*

It is obvious that one could easily be trapped into using invalid measures to test H_1, H_2, and H_3. All you have to do is ask questions that refer to the aggressive violence that is absent in one version of a program or questions that refer to color-specific aspects of peripheral material. Such questions would make it quite difficult for subjects exposed to a low aggressive violence version of a program or a black and white presentation to score as well as subjects in the other experimental groups. However, the intent of the hypotheses was to predict that *more material of a certain type, that was equally available to all subjects, would be learned under certain circumstances.*

The fourth hypothesis tested in this study involved an expected interaction between high aggressive violence and color format. It was expected that the two factors would combine to produce effects even stronger than they produced alone.

H_4: *A color presentation of high aggressive violence will cause more recall of peripheral-visual material, relative to central-visual material, than would have been expected from the simple effects of color and aggressive violence.*

Age of young viewers

Two predictions were made for the differences in learning scores by different age groups:

H_5: *Older children will recall more central material.*

H_6: *Younger children will recall more peripheral material in relation to central material recalled.*

H_5 is a simple statement of the obvious expectation that older children remember more of what they have seen. H_6 reflects the finding that older children tend to focus more on the essential material than do younger children. It does not predict that more peripheral content will be recalled by younger viewers. Rather, it predicts that the balance between central and peripheral will show a relative emphasis on peripheral material among younger children.

The effects of delayed testing and other interactions

Most of the research on media effects has used tests administered almost immediately after the presentation of the message. When this study was designed, the question of whether the effects would be stable over time was raised. As a result, it was decided to test some subjects for recall immediately after the presentation and other subjects two weeks after the presentation. The obvious prediction that could be made was that all types of recall would be poorer after a time delay. This did not seem worthy of elevation to the position of a formal hypothesis in the design. All other predictions that were made before the data were collected were more in the nature of guesses based on intuitive feelings than hypotheses derived from a theoretical base. Thus, while the time delay was an interesting and potentially important factor, the project had to be considered a pilot study that might provide data. Such data could then lead to further insights, hypotheses, and research.

Similarly, there are numerous potential interaction effects between violence, color, age of viewer, and central/peripheral distinction. Since no data were available to indicate what these might be, we decided not to present intuitive speculations as hypotheses. Rather, the study was designed to test a small number of hypotheses and to provide a description of phenomena about which there was no clear prior evidence.

EXPERIMENTAL METHOD

The experiment tested for effects in 24 different experimental conditions. These were determined by three age groups, two formats, two levels of violence, and two testing dates. Figure 1 presents a representation of the design.

			Grade					
			4th TEST		6th TEST		9th TEST	
		AGGRESSIVE VIOLENCE	Immediate	Delay	Immediate	Delay	Immediate	Delay
Format	Color	High	1	2	9	10	17	18
		Low	3	4	11	12	19	20
	B/W	High	5	6	13	14	21	22
		Low	7	8	15	16	23	24

Figure 1: Schematic representation of study design

This design is four-dimensional. Analyses of the data, therefore, required the use of a four-way analysis of variance for each dependent variable considered.

Subjects

Subjects were boys drawn from within the same school district. Males were used because the addition of female subjects would have doubled the number of subjects required and because the primary concern in studies of violence to date has been with effects among male subjects. The school district in which the study was made was middle class to upper-middle class, suburban, and predominantly white.

The three age groups studied were fourth graders (9-10 years old), sixth graders (11-12 years old), and ninth graders (14-15 years old). The fourth grade was chosen because it was the youngest group from which we expected sufficient reading ability to deal with a written questionnaire. The ninth grade was chosen because it had been shown to be on a plateau beyond which age is not a significant factor in cognitive development. The sixth grade was chosen to fit between the two.

The ninth-grade boys were all taken from the same high school. The sixth-grade boys were all taken from the same middle school. The fourth-grade boys were distributed among three elementary schools. Students from the elementary schools would be expected to enter the middle school studied, and students from the middle school would be expected to enter the high school studied. Thus, the grade level of the subjects was varied while the general demographic characteristics of their homes was assumed to be constant.

Class lists were obtained for each grade level. A random number table was then used to assign 12 boys to each experimental cell. (Only 93 boys were available in the sixth grade, so some cells had 11 subjects.) Each cell was then further divided into two runs with five or six boys per run. One of the elementary schools provided half of the fourth grade subjects; they were randomly assigned to one run in each of the eight conditions. The other two elementary schools each provided subjects for one run of four experimental conditions.

The decision to choose as many as 12 boys per cell had been taken with the hope that a minimum of ten boys per condition would be available for the delayed test. In the fourth and ninth grades, alternates were also chosen in case the number per condition in the original presentation of the program fell below eleven. Two hundred seventy-eight boys were shown the program initially. Of these, 33 were unavailable at the time of the delayed test. However, a callback test, the morning after the first delayed test, insured ten subjects per cell. Fourteen of the experimental cells contained more than ten subjects (total = 260 for all 24 cells), and each condition was reduced to ten subjects through random elimination. This was done so that the complex analysis of variance procedure could be simplified through the use of equal size cells. A cursory evaluation of cell means before and after subjects had been eliminated indicated that no differences in average scores had resulted from this procedure.

In sum, the data in this study came from ten middle-class boys from a single school district in each of 24 experimental conditions. There were 80 boys in each age group. There were 120 boys in each color format. One hundred twenty boys saw each of two levels of aggressive violence. There were 120 boys tested immediately after the program and 120 boys tested after a two week delay.

The manipulation of aggressive violence

After a considerable search, a stimulus program was obtained which met the following criteria: half-hour length, originally broadcast in color, broadcast nationally more than five years prior to the date of the study, not rerun in the area of the study, containing aggressive violence that could be removed. Two versions of the program were then created by eliminating between two and three minutes of aggressive violence from one version (Low Violence) and eliminating between two and three minutes of nonviolent content from the other version (High Violence). Appendix A gives a summary of the two programs. The editing attempted to be unobtrusive and to balance the portions of the program from which the nonviolence and the violence were removed. I believe we were quite successful at both tasks. Both the Low and High Violence versions run quite smoothly with no indication that anything has been removed.

In fact, very little had been removed. However, the High Violence version contained a struggle, a slap, a glass thrown against a wall, a point-blank murder (complete with bullet holes), a gun battle, and a violent fistfight. In the Low Violence version the murder was implied and then proven, and the gunbattle-fistfight became a chase with a one-karate chop ending. Both versions of the program ran just over twenty minutes including titles. (Commercials and the short wrapup scene were not included.)

The color variable

Several copies of both versions of the program were recorded in color on one-inch videotape. The project used two similar videotape systems, each of which could play the tapes in color. One system was an Ampex 7500C videotape machine, a color corrector unit, and an RCA 23-inch color television monitor. The color system could be made to play the same tape in black and white with the simple turn of a knob on the color corrector. The second system was an Ampex 4900C color videotape player and another RCA 23-inch color monitor. This system could be made to play color tapes in black and white with a change of the position of an output cable. So the manipulation of the color variable was very simple. All subjects saw the same videotapes on the same type of television sets. The experimenter merely made a minor technical adjustment to play the tapes in color or black and white.

Immediate and delayed tests

The design of the study was a bit more complex than Figure 1 might indicate. It was decided that some opinion items should be given to the group of boys that was not tested for recall immediately after the program. This was done for two reasons. First, there was a need to explain why we were showing these boys the program. Second, the data were of interest and might provide important insights into the other effects studied. Thus, opinion items were added to the design.

The subjects in all of the odd-numbered cells in Figure 1 were tested for recall immediately after the presentation. The subjects in the even-numbered cells in Figure 1 were asked for their opinions of the program immediately after the presentation. All subjects were asked for both opinions and recall in a delayed test that was given two weeks after presentation. This design allows for immediate/delayed comparisons between independent samples for recall (odd-numbered cells for immediate, even-numbered cells for delay) and for opinions (even-numbered cells for immediate, odd-numbered cells for delay). As a result, the effects of prior testing and familiarity with the test items were minimized. The design also allows a test for the effects of prior testing (comparison of two-week delayed scores between subjects tested immediately and subjects not tested immediately); however, this is a methodological question that does not need to be covered in this report.

The instrument

Immediately after exposure to the stimulus presentation, subjects were given one of two testing instruments. Those who were tested on learning were asked about their familiarity with the program, given 31 multiple choice information questions, asked about their home viewing habits, and then given the picture ranking procedure (see below). Those who were tested for opinions were given 15 opinion items instead of the information items. After a two-week delay period, all subjects were given 31 information questions and 15 opinion questions as well as the picture ranking task.

Recall items. Since there was concern over the relationships between different types of learning, great care was taken in the construction of recall items. Language had to be clear and simple, so that the items could be understood by nine-year-olds while 15-year-olds would not think they were silly or too simple. Care was taken to devise items testing for information that was equally available in both the High and Low Violence versions. Items also had to meet criteria that had been set for dividing the questions into four mutually exclusive categories.

Two pretests were run to help obtain a valid set of central and peripheral items. First, 52 items were written regarding all aspects of the content common to both versions of the program. Then 36 judges (college undergraduates) saw either of the two versions and answered the questions. They indicated whether they thought a question was central ("has to do with the plot or main action of the story, or concerns any important part of the program") or peripheral ("has little to do with the plot or action, or has nothing to do with the main activity") or neither central nor peripheral. Items that were unanimously placed in the central or peripheral category by the three researchers responsible for the items and that were placed in the same category by more than two-thirds of the judges were retained. Others were thrown out or rewritten. At this point, a second division of items was made. Those which could be answered by seeing the video portion of the program and could not be answered if only the sound track was played were called *visual* items. Those which could be answered after listening to the sound track and could not be answered from the visual track were called *auditory* items. For some items this distinction was not perfect, but the primary source of information was obvious; these were retained. Items were discarded if there was ambiguity about the primary source of the information.

A set of 36 items was selected for the second pretest. Fifty judges from a class in "Communication Effects" were asked to make central/peripheral judgments. Items that were placed in the appropriate category by 80 percent or more of the judges were retained. Twenty-seven items were acceptable under this criterion. Three items which almost met this criterion were rewritten so that they were acceptable to all three researchers working on the questions. The result of the pretesting was the set of 30 items presented in Appendix B. There were 15 central items —eight central-auditory (CA) and seven central-visual (CV)—and 15 peripheral items—eight peripheral-auditory (PA) and seven peripheral-visual (PV). An additional item ("Before the killer murders Egan he says. . ." Answer: Nothing) was inserted to see if there was a tendency to imagine that something had happened when it actually had not appeared in the program. The 31 information items were presented in random order. For the delayed test, a different random order was used and the alternative choices were rearranged.

Opinion items. The 15 opinion items all presented subjects with four alternatives. The boys were instructed to check the "best answer." Nine items referred to the show, one referred to police programs in general, and five referred to characteristics in the program. Two items (Was the show "not very," "pretty," "very," or "extremely" violent? Was the show "not very," "a little," "pretty," or "very" mean?) served as an index of perceived violence. There was no assumption that the raw values from this scale could be interpreted; however, differences across experimental conditions were taken as indicators of the effects of manipulations. Four items were used as an indication of how much the program was liked (Was it a good or bad thing to watch? Was it a show like you really like to see or don't like to see? Was it a wonderful or terrible show? Did you enjoy it very much or not at all?). This scale was also used as a comparison across experimental conditions.

Other items. All subjects were asked if they had ever seen any episode from this series and if they had seen this particular episode before. Although older boys were more likely to recognize the series, only four ninth-grade, four sixth-grade, and six fourth-grade boys said they had seen this episode before. At the end of the questionnaire, subjects were asked to indicate if they watched television "a lot," "sometimes," "very little," or "almost never." A final question asked if they usually watched a color or a black and white set.

Picture ranking task. After all subjects in a run had completed the printed questionnaire, each one was given a packet containing 12 4" x 5" photographs. The photographs were of scenes common to both versions of the program. All subjects were given the same 12 photographs, printed in black and white and presented in the same random order. Subjects were then asked to put the photographs into the order in which the scenes had appeared in the program. This task was taken to be an indication of the subjects' ability to recall central visual material.

Procedures

Testing was done over a ten-day period. Three consecutive days were required in the high school, two consecutive days were required in the middle school, two days were required in one elementary school, and a half-day was required in each of the other two elementary schools.

Subjects were randomly assigned to one of 16 runs (two for each condition) within each age group. They were escorted to small rooms within their schools in groups of five or six. All testing took place during the school day, and all boys were released from 45 minutes of school time by permission of their teachers.

Two experimental rooms were run simultaneously in each school. All rooms used were fairly similar. They were quiet and comfortable. Subjects had a table or desk on which to write. The television monitor was in front of the room, and the videotape machine was out of the line of sight to the monitor.

In each of the two rooms there was a female experimenter (E_1) who was in charge and a male experimenter (E_2) who ran the videotape equipment. Both people who served as E_1 were attractive and articulate young women. The assignment of experimental conditions was counterbalanced so that each E_1 was in charge of one run of each condition and so the two experimental rooms were never running the same experimental condition. The two runs of each condition were also placed so that there was an optimum balance of time of day and day of test.

The subjects entered the experimental room, where E_1 would ask them to be seated and check their names against an attendance sheet. There were no papers, pencils, or questionnaires visible to subjects at this point. E_1 would then say, "Now we'd like you to watch this television program." She would turn off the lights and E_2 would start the videotape.

At the conclusion of the program E_2 would turn off the machine—to minimize noise in the room, and E_1 would distribute questionnaires and pencils. She then said, "We'd like you to answer a few questions about the program you just saw. This isn't a test, so just give the best answers you can and don't worry about how anybody else is doing. If you have any questions just raise your hand, and I'll come to you and help." The procedure worked very well; a few fourth graders had questions about a word they did not understand or could not read, but this number was small, and their problems were easily answered.

When every subject had finished with the printed form, E_1 distributed the packets of photographs and asked the subjects to place them in the order in which they appeared in the program. On completion of this task, the group was asked to keep the whole thing a secret for a few days "because it wouldn't be fair if the other people we ask these questions know what to expect. You didn't, did you?" The group was then dismissed, and E_1 and E_2 copied the rank order of the pictures on the bottom of the last page of each subject's questionnaire.

Pictures were put back in the preset order, the videotape was cued for the next run, the room was tidied, and the experimenters were ready for the next group of subjects.

Thirteen to 15 days after the first test, the second questionnaire was administered. Subjects had not been told about the delayed test and did not expect to be asked about the program they had seen two weeks earlier. Fourth-grade boys were tested by class; all boys in each classroom received the delayed test at one time. Ninth-grade boys were also tested by class; they had been drawn from social studies classes. Sixth grade boys were all tested at one time in one room; about eighty boys were brought to the school cafeteria. The delayed test required about 25 minutes. It included all questions on information and opinions as well as a repeat of the picture ranking task.

For each grade the delayed test was completed in less than one day per school. If subjects had been absent, an attempt was made to administer the delayed test to them individually on the following morning. Fifteen subjects received the delayed test on this type of callback.

RESULTS

The results of the experiment were essentially analyzed by way of the four-way analyses of variance presented in Appendix C. This section of the report groups the findings in terms of the independent variables studied. The thirteen major dependent variables were:

1. *Total learning*—the total number of items for which a subject gave a correct answer. The maximum possible score was 30. The mean score for all subjects on both tests was 20.55.
2. *Peripheral*—the total number of peripheral items correctly answered. Maximum possible was 15. Mean was 9.15.
3. *Peripheral-auditory*—the total correct in this category. Maximum was 8. Mean was 4.63.

4. *Peripheral-visual*—the total correct in this category. Maximum was 7. Mean was 4.52.
5. *Central*—the total number of central items correctly answered. Maximum possible was 15. Mean was 11.41.
6. *Central-auditory*—the total correct in this category. Maximum was 8. Mean was 5.52.
7. *Central-visual*—the total correct in this category. Maximum was 7. Mean was 5.89.
8. *Peripheral-central difference*—the value of total peripheral minus total central items. Since the central scores were higher than the peripheral scores, the mean score had a negative value (−2.25). Higher values of this variable indicate better recall of peripheral material in relation to central material.
9. *PA-CA difference*—total peripheral auditory minus total central auditory items. This is a subcategory of peripheral-central difference scores that considers only material from the sound track. Mean score was −.89.
10. *PV-CV difference*—total peripheral visual minus total central visual items. Mean score was −1.37.
11. *Rank.* For each subject, the numerator of a rank order correlation coefficient between his ranking of photographs and the true rank order. The denominator for this coefficient is a constant (64) for a constant number (12) of items, so it was ignored. Scores could potentially range from −64 to +64. Mean score was 46.11.
12. *Violence index.* A factor analysis indicated that the two items intended to indicate perceived violence had formed a clear factor from among the fifteen opinion items. Scores on the items were added to give an index ranging from 2 to 8. The overall mean was 4.59. Higher scores indicated more perceived violence.
13. *Liking index.* The four items intended as an index of liking of the program also formed a clear factor among the opinion items. The index value could range from 4 to 16. The overall mean was 7.37. Lower scores indicated a more positive evaluation of the program.

The effects of level of violence

Table 1 is a summary of the effects of the two levels of aggressive violence on the different variables studied. It is clear that *subjects considered the High Violence presentation significantly more violent than the Low Violence presentation* ($p < .0005$). Thus the manipulation of the violence level was successful.

There was no support for either H_1 or H_2. The evidence suggests that there is no significant difference in learning between high and low violence programs. *In one case (peripheral material), significantly more was*

Table 1: Effects of level of violence

	Means by violence levels				
VARIABLE	High	Low	F	P	Interaction
Total learning (30 items)	20.32	20.80	1.11	x	grade
Peripheral (15 items)	8.91	9.39	3.93	.05	grade x time
Per-Auditory (8 items)	4.51	4.74	2.10	x	grade x time
Per-Visual (7 items)	4.39	4.65	2.88	x	grade x color
Central (15 items)	11.41	11.41	x	x	grade
Cent-Auditory (8 items)	5.48	5.56	x	x	grade
Cent-Visual (7 items)	5.92	5.85	x	x	
Peripheral-central diff.	−2.50	−2.02	2.61	x	
PA-CA diff.	−0.97	−0.82	x	x	color x grade x time
PV - CV diff.	−1.53	−1.20	3.66	x	color color x grade color x time
Rank (maximum = 64)	46.78	45.43	x	x	
Violence index (high=high)	4.96	4.22	15.07	.0005	
Liking index (low=better)	7.48	7.26	x	x	

recalled from the Low Violence version. The scores on peripheral-central differences were not significant and were in the opposite direction from the prediction of H_2.

In general, the level of violence did not seem to have a clear effect on learning. However, Table 1 indicates that there were many significant interaction effects; these will be considered in a later section of this paper. Violence does not seem to be a simple cause of learning phenomena. It does seem to interact with other factors in the viewing situation.

Two notes of caution should be introduced when the data from two violence levels are interpreted. First, the two programs were not identical. Although the editing was minimal and subtle, it cannot be claimed that subjects saw identical programs. The questions used were limited to content that was common to both programs; however, there is always the possibility that cues in one version of the program made some questions easier for subjects who saw that version. The evidence (e.g., mean scores for total learning) does not support this type of difference. Yet it can never logically be excluded. This type of cautionary note—that high and low violence stimuli differ on many dimensions other than violence—should also be applied to studies of behavioral effects due to "violent" programs.

The second problem in this study was that it was impossible to meaningfully compare amount of *violent material* recalled. You simply cannot recall violence that is not present in a Low Violence presentation, so it was not possible to test differences. Thus, none of the learning items are about recalled violence. Perhaps the study was flawed by lack of questions (about violent action) that might have been given only to the

High Violence subjects, with analyses by color presentation and age group. Such a procedure should prove fruitful in further research.

The effects of color presentation

The evidence summarized in Table 2 indicates that H_3 was supported. *Color presentation resulted in a relatively better recall rate for peripheral-visual minus central-visual material.* This effect was demonstrated at an extremely high level ($p < .005$) of statistical significance. In part, it can be interpreted in light of the only other main effect of color format. *The group that viewed the programs in black and white recalled more central-visual material than the groups that saw a color presentation.* This accounts, to a large extent, for the difference in PV-CV scores. It also indicates that a black and white presentation may be more effective in presenting the central aspects of a visual message. Unfortunately, the picture ranking task—which was expected to be another indicator of central-visual recall—did not produce significant differences. In retrospect, it seems that this procedure was not a clear test of central material only.

Table 2: Effects of color presentation

	Means by format				
VARIABLE	Color	B/W	F	P	Interactions
Total learning (30 items)	20.39	20.73	x	x	
Peripheral (15 items)	9.12	9.18	x	x	
Per-Auditory (8 items)	4.57	4.69	x	x	
Per-Visual (7 items)	4.55	4.49	x	x	violence grade violence x grade
Central (15 items)	11.28	11.54	x	x	
Cent-Auditory (8 items)	5.61	5.43	x	x	
Cent-Visual (7 items)	5.67	6.10	7.83	.01	
Peripheral-central diff.	−2.16	−2.36	x	x	
PA-CA diff.	−1.04	−0.74	1.78	x	violence x grade x time
PV-CV diff.	−1.12	−1.62	8.28	.005	violence violence x grade violence x time
Rank (maximum = 64)	46.25	45.97	x	x	
Violence index (high=more)	4.47	4.71	1.60	x	
Liking index (low-better)	7.24	7.50	x	x	

H_4 predicted an interaction between color and level of violence on PV-CV difference scores. Table 2 indicates that this interaction was statistically significant. (See Table C-12, Appendix C.) Table 3 presents the mean scores by violence level and format condition for PV-CV difference scores and peripheral-visual scores (interaction also statistically significant).

Table 3: Mean scores by violence level and format condition

A. PV-CV difference scores

	Low violence	High violence
Color	−1.166	−1.066
B/W	−1.233	−2.000

B. Peripheral visual scores

	Low violence	High violence
Color	4.416	4.683
B/W	4.883	4.100

The data from Table 3 are presented graphically in Figure 2 and Figure 3. Figure 2 indicates partial support for the prediction of H_4. *The color format produced higher PV-CV difference scores than the black and white format for the High Violence presentation. For the Low Violence presentation, the scores are similar.* An unusual aspect of this phenomenon is the indication that B/W scores were much lower in the High Violence condition than in the Low Violence condition. The effect is partially attributable to the finding, shown in Figure 3, that the peripheral-visual scores for the B/W-High Violence group tended to be much lower than those for the B/W-Low Violence group and the Color-High Violence group.

Table 4: Effects of different grade levels

	Means by Grade levels					
VARIABLE	4th	6th	9th	F	P	Inter-actions
Total learning (30 items)	17.79	21.30	22.59	39.25	.0005	violence
Peripheral (15 items)	7.98	9.39	10.09	25.97	.0005	violence x time
Per-Auditory (8 items)	4.01	4.70	5.18	18.93	.0005	violence x time
Per-Visual (7 items)	3.96	4.69	4.91	14.19	.0005	color color x violence
Central (15 items)	9.81	11.91	12.50	29.53	.0005	violence
Cent-Auditory (8 items)	4.40	5.79	6.38	33.97	.0005	violence
Cent-Visual (7 items)	5.41	6.13	6.13	9.06	.0005	
Peripheral-central diff.	−1.83	−2.52	−2.41	2.03	x	
PA-CA diff.	−0.39	−1.09	−1.20	5.11	.01	violence x color x time
PV-CV diff.	−1.45	−1.43	−1.22	x	x	violence x time color x violence
Rank (maximum = 64)	38.05	48.73	51.55	21.00	.0005	
Violence index (high=high)	4.83	4.78	4.16	4.97	.01	
Liking index (low=better)	6.08	7.71	8.32	20.14	.0005	

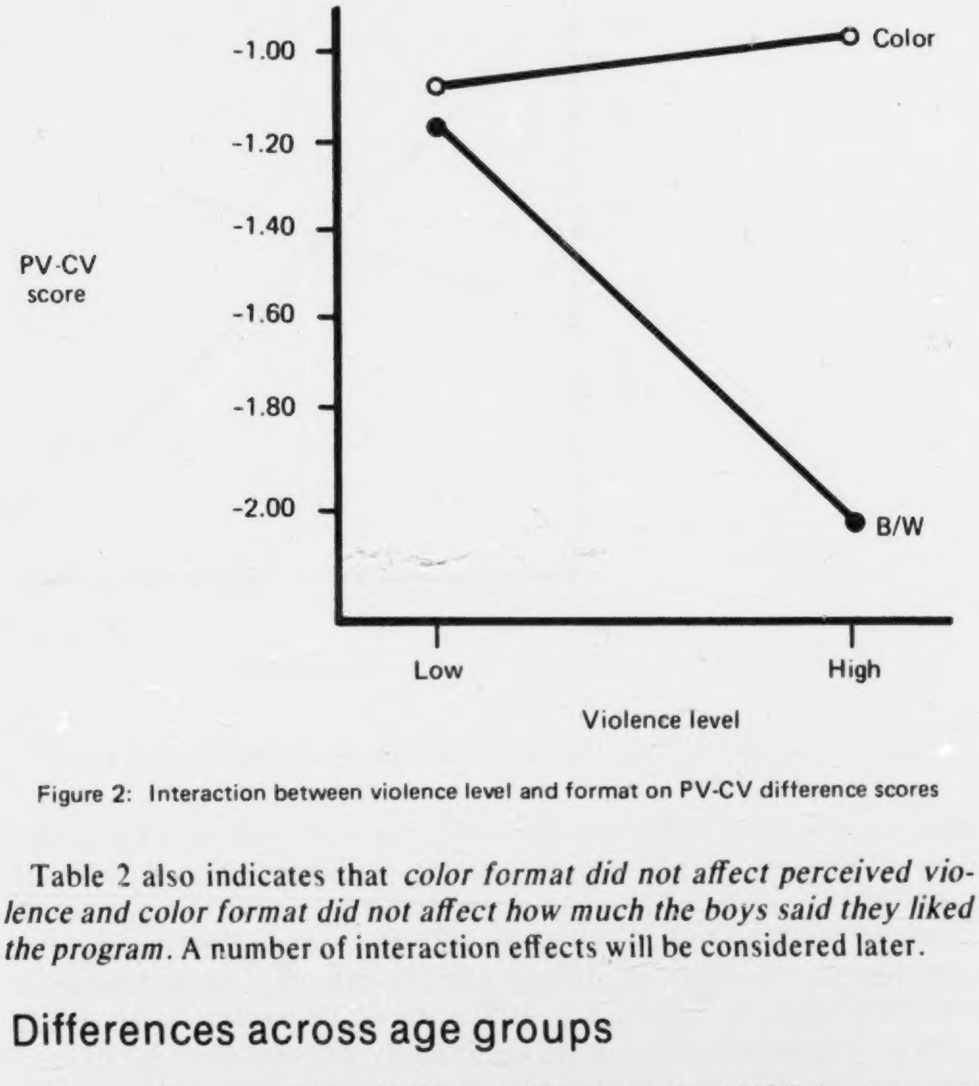


Figure 2: Interaction between violence level and format on PV-CV difference scores

Table 2 also indicates that *color format did not affect perceived violence and color format did not affect how much the boys said they liked the program*. A number of interaction effects will be considered later.

Differences across age groups

There were extremely significant differences ($p<.0005$) between age groups on all seven learning scores and the picture ranking task. Table 4 indicates that sixth graders tended to do better than fourth graders and that ninth graders tended to do better than sixth graders. Figure 4 presents a graphic representation of the data for the different recall scores over the three grades tested. It can be seen that the *increases from fourth to sixth grade were much sharper than the increases from sixth to ninth grade*. The monotonic increase in recall of central material supports the prediction of H_5. Peripheral recall also increased monotonically over the three grades; this fails to reaffirm the curvilinear relationships reported by other researchers.

The main effect of grade in the four-way analysis of variance of peripheral-central difference scores (Appendix C, Table C-10) did not reach a level of statistical significance. Comparisons of separate grades indicated that the difference score for the fourth grade was significantly

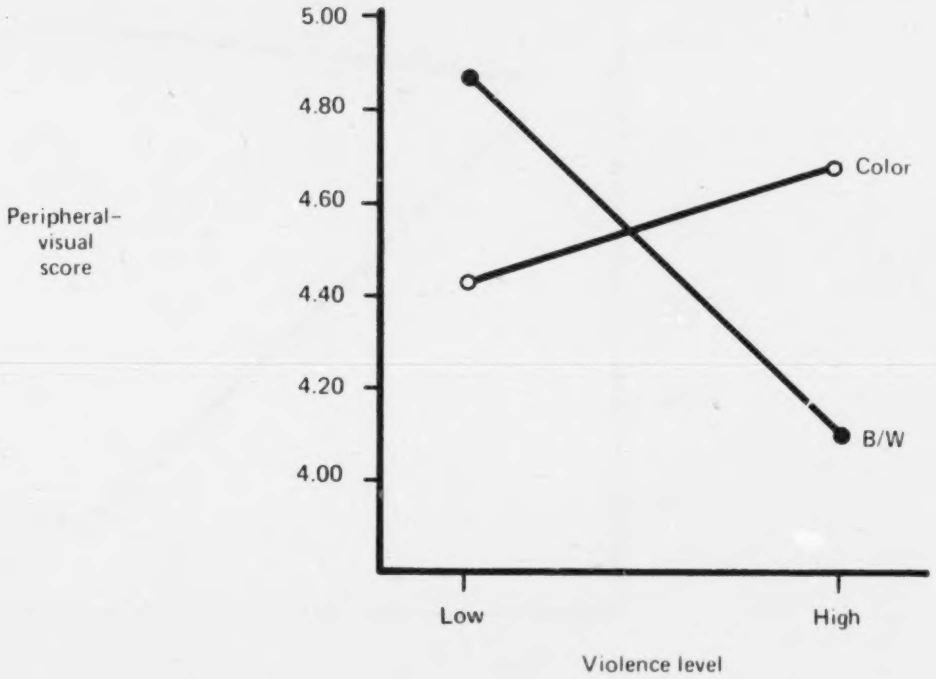


Figure 3: Interaction between violence level and format on peripheral-visual scores

higher (t = 1.89, $p < .05$) than the difference score for the sixth grade. (Lower negative values—higher scores—indicated higher ratio of peripheral to central.) The ninth-grade difference score was not significantly different from the sixth-grade difference score. This finding provides partial support for H_6. It seems that *fourth-grade boys recall more peripheral material, relative to central material, than older boys.*

PA-CA difference scores were significantly different ($p < .01$) over all three grades, providing further partial support for the hypothesis. *For material presented by the sound track older boys tended to focus more on central activity in the program.* This finding is not replicated for material presented primarily by the video portion of the program.

Results from the liking index demonstrated that *younger boys were significantly more favorable ($p < .0005$) in their evaluations of both versions of the program than older boys.* Sixth-grade boys rated the High Violence and Low Violence programs as equal on the liking index. Ninth-grade boys did the same. *Fourth-grade boys indicated that they liked the Low Violence version significantly better (t= 2.04, $p < .025$) than the High Violence version.* The finding that younger boys said they liked the program better than older boys cannot be attributed to a simple tendency for young boys to give more positive responses, since it was also found that *for each violence level, younger boys were significantly higher ($p < .01$) in the amount of violence perceived in the program.* These significant differences on the perceived violence index and the lik-

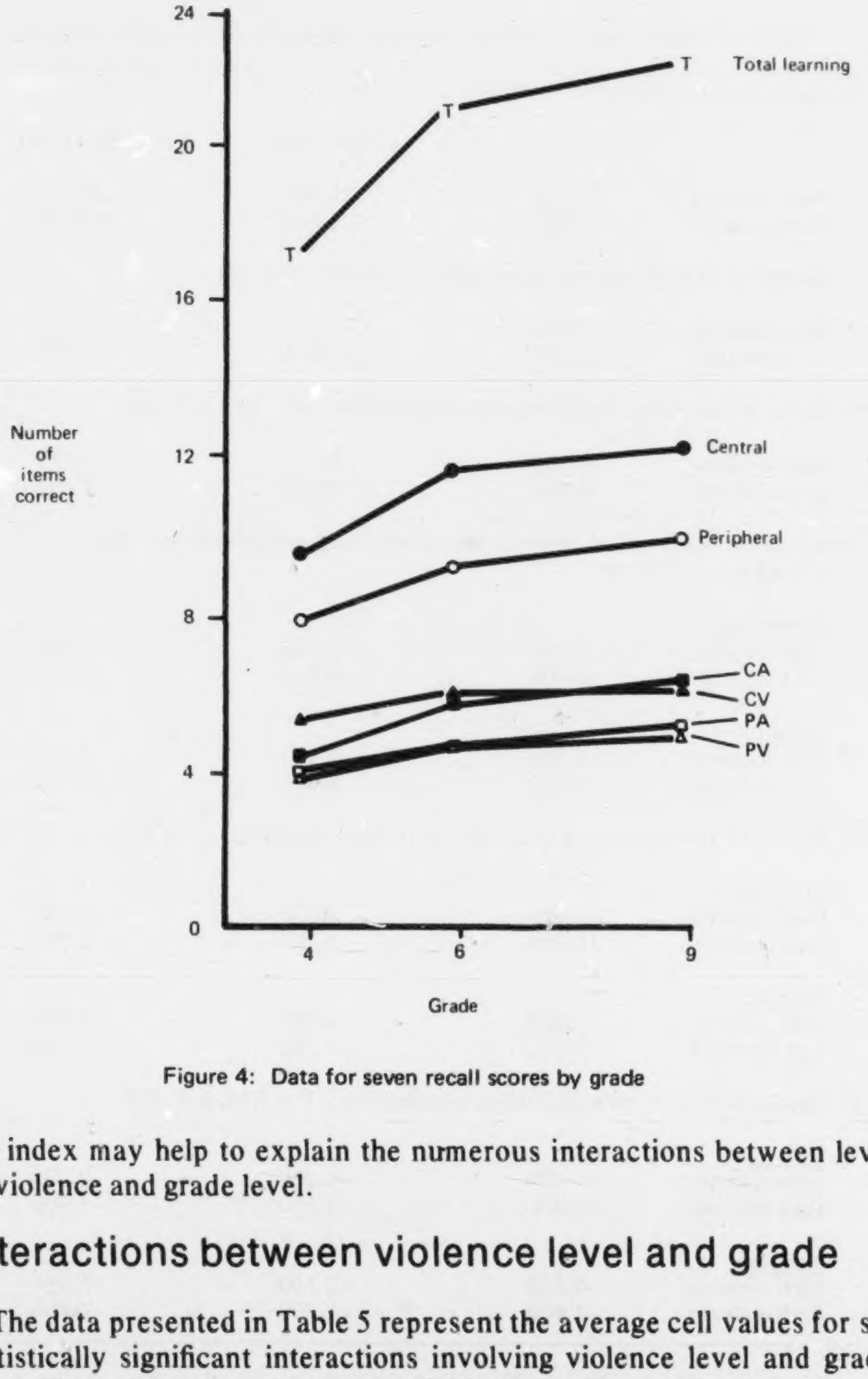


Figure 4: Data for seven recall scores by grade

ing index may help to explain the numerous interactions between level of violence and grade level.

Interactions between violence level and grade

The data presented in Table 5 represent the average cell values for six statistically significant interactions involving violence level and grade level. These data are represented graphically by Figures 5 - 10.

It can be seen that the effects of High and Low Violence presentations were reversed from the fourth to the ninth grade, with the sixth graders doing almost equivalently well for both levels of violence.

Fourth graders learned more from the Low Violence version of the program.

Table 5: Mean scores for interaction effects between level of violence and grade

A. Scores on total learning (Interaction: F = 3.94, $p < .02$)

	4th Grade	6th Grade	9th Grade
High violence	16.750	21.075	23.125
Low violence	18.825	21.525	22.050

B. Scores for central material (Interaction: F = 3.65, $p < .03$)

	4th Grade	6th Grade	9th Grade
High violence	9.350	11.850	13.025
Low violence	10.275	11.975	11.975

C. Scores on central-auditory material (Interaction: F = 3.11, $p < .05$)

	4th Grade	6th Grade	9th Grade
High violence	4.100	5.675	6.675
Low violence	4.700	5.900	6.075

D. Scores for peripheral material (3-way interaction, including time of test: F = 4.68, $p < .01$)

	4th Grade	6th Grade	9th Grade
Immediate			
High violence	7.700	10.300	11.650
Low violence	9.700	10.650	10.650
Delayed			
High violence	7.100	8.150	8.555
Low violence	7.400	8.450	9.500

E. Scores for peripheral auditory material (3-way interaction: F=4.78, $p < .01$)

	4th Grade	6th Grade	9th Grade
Immediate			
High violence	3.700	5.050	5.650
Low violence	4.950	5.250	5.450
Delayed			
High violence	3.850	4.350	4.500
Low violence	3.550	4.150	5.100

F. Scores for PV-CV difference (3-way interaction: F = 3.15, $p < .05$)

	4th Grade	6th Grade	9th Grade
Immediate			
High violence	−1.700	−1.200	−0.900
Low violence	−1.050	−1.250	−1.400
Delayed			
High violence	−1.550	−2.100	−1.750
Low violence	−1.500	−1.200	−0.800

Ninth graders learned more from the High Violence version of the program.

This interaction effect was significant for total learning, central material, and central-auditory material. (See Figures 5-7.) Figures 8-10 indicate that the same crossover appeared, *for those subjects who were tested immediately after the presentation,* on scores for peripheral material, peripheral-auditory material, and PV-CV difference scores. The circles, solid and hollow, in Figures 8-10 represent data collected immediately after the presentation. They show the same interaction as Figures 5-7. The hollow and solid squares represent data collected after a two-week delay. For these three types of recall, the patterns were not the same in the two time periods.

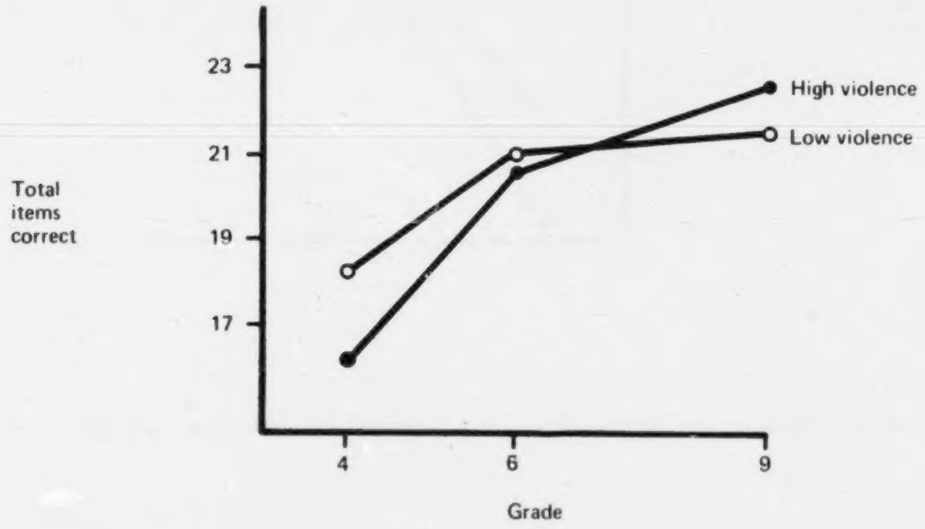


Figure 5: Interaction between violence level and grade on total learning scores

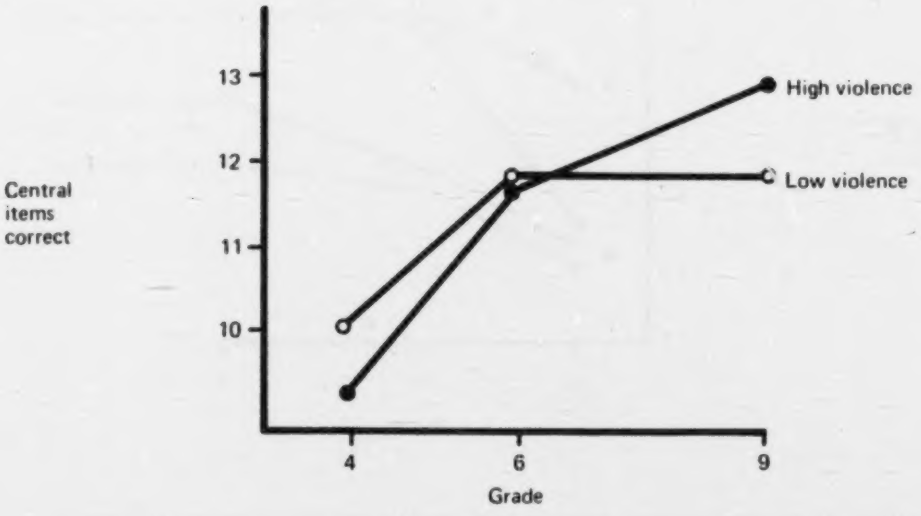


Figure 6: Interaction between violence level and grade on central learning scores

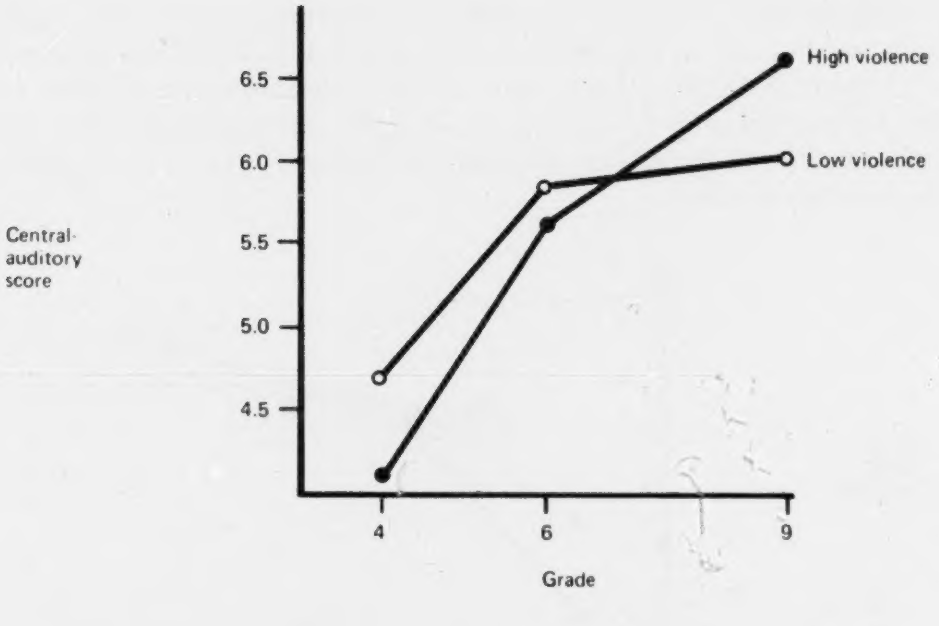


Figure 7: Interaction between violence levels and grade central-auditory scores

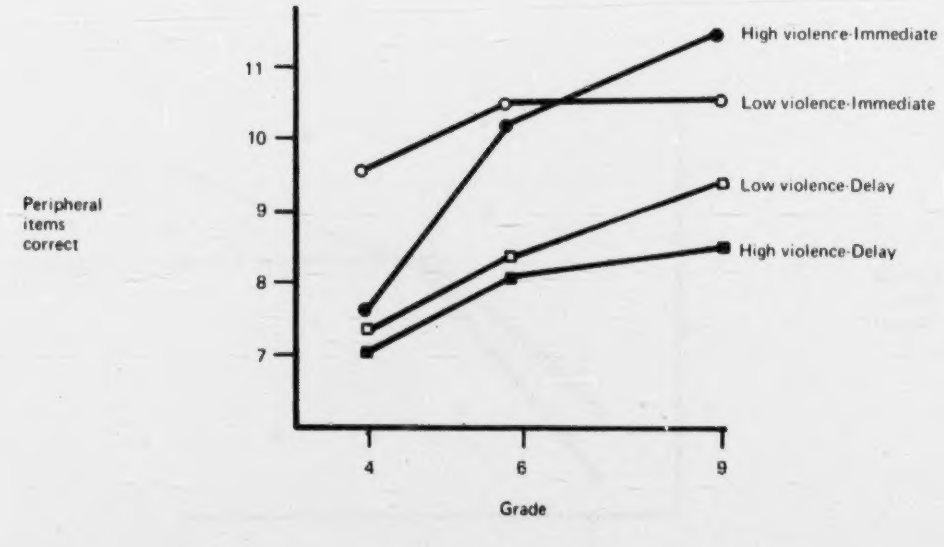


Figure 8: Interaction between violence level, grade, and time of test for peripheral learning scores

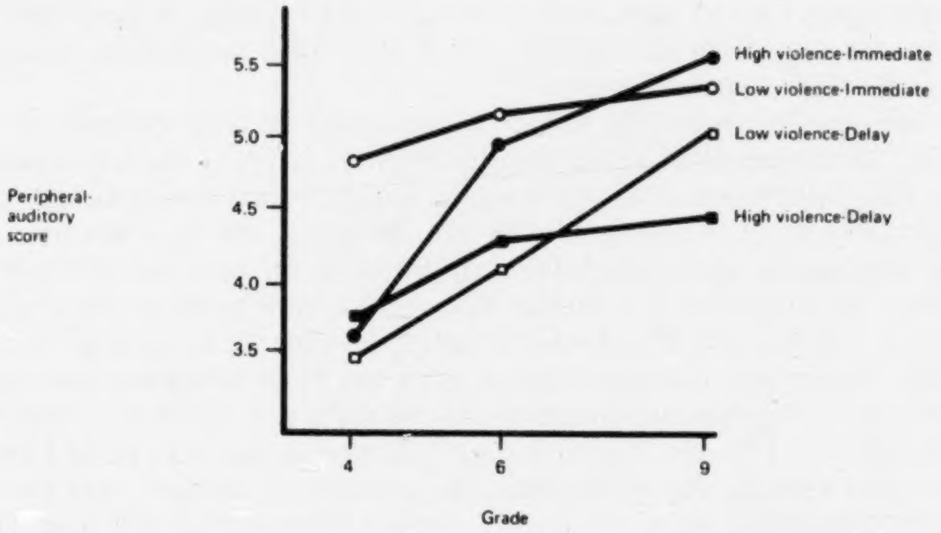


Figure 9: Interaction between violence level, grade, and time of test for peripheral-auditory scores

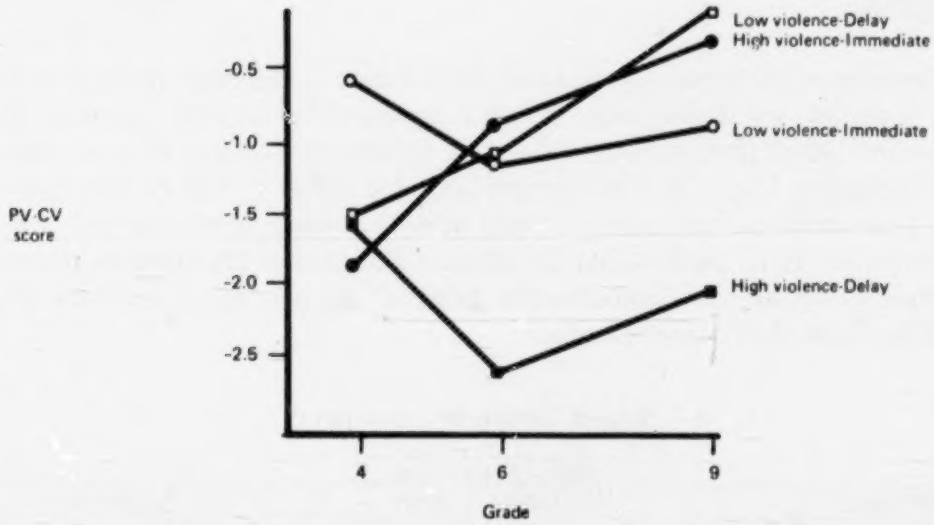


Figure 10: Interaction between violence level, grade, and time of test for PV-CV difference scores

When these interactions are interpreted in light of the results on the violence index and the liking index, we find that the ninth-grade boys, who perceived least violence and liked the programs least, learned more from the High Violence version of the program. Sixth-grade boys, who were midway between the other two grades on perceived violence and program evaluations, did equally well on both levels of violence. Fourth-grade boys, who perceived higher levels of violence and liked the programs better than boys in the other two grades, learned more from the Low Violence version—which they liked significantly better than the High Violence version.

Since the fourth-grade boys learned more from the Low Violence version, which they liked better than the High Violence version, it seems possible that there is a tendency to avert attention from programs high in aggressive violence among boys in that age group. On the other hand, the ninth-grade boys may have been drawn to the program that contained more aggressive violence. They found both versions relatively low in violence, and the additional aggressive violence in the High Violence version may have made the program more interesting and exciting for them. They indicated that they thought there was significantly more "action" ($t= 1.79$, $p< .05$) in the High Violence version than in the Low Violence version. For adolescents, the question of "action" may be a more relevant indicator of what boys see in a program than responses to "good," "like to see," "violent," or "mean."

Differences between immediate and delayed tests

It came as no surprise to find that *for all recall categories subjects who were tested two weeks after seeing a presentation recalled significantly less ($p< .0005$) than subjects who were tested immediately after seeing a presentation.* These findings are presented in Table 6. The picture ranking task produced equivalent scores at both testing times; but this may have been due to the fact that all subjects had ranked the pictures immediately after the presentation—the delayed test was invalidated by the anchoring of the immediate test.

Table 6: Effects of time of test

	Means for each time				
VARIABLE	Immediate	Delay	F	P	Interactions
Total learning (30 items)	22.550	18.566	75.67	.0005	
Peripheral (15 items)	10.108	8.191	61.78	.0005	violence x grade
Per-Auditory (8 items)	5.008	4.250	23.89	.0005	violence x grade
Per-Visual (7 items)	5.100	3.941	51.94	.0005	
Central (15 items)	12.441	10.375	47.94	.0005	
Cent-Auditory (8 items)	6.091	4.950	32.29	.0005	
Cent-Visual (7 items)	6.350	5.425	34.34	.0005	
Peripheral-central diff.	–2.333	–2.183	x	x	
PA-CA diff.	-1.083	-0.700	2.91	x	violence x color x grade
PV-CV diff.	-1.250	-1.483	1.80	x	violence x grade
Rank (maximum = 64)	45.866	46.350	x	x	
Violence index (high=high)	4.375	4.800	4.95	.03	
Liking index (low=better)	7.250	7.491	x	x	

The group that gave opinions of the programs after a two-week delay said that they thought the programs were significantly more violent than the group that gave opinions immediately after the presentation. Table 7 presents the violence index scores for the two presentations at the two time periods. The difference between immediate and delayed ratings of the Low Violence version is small and does not approach statistical significance. The difference between ratings of the High Violence version is larger and is statistically significant. This indicates that *after a two-week period the High Violence version was rated significantly more violent than it had been rated immediately after presentation, while the rating of the Low Violence version did not change significantly.*

Table 7: Scores on index of perceived violence
by level of violence and time of test

	Immediate	Delayed	t
High violence	4.633	5.283	2.20
Low violence	4.116	4.316	.80 (p. < .02)

One other critical effect occurred over time and was not revealed by the four-way analyses of variance. Table 1 indicates that the differences between High Violence and Low Violence groups on PV-CV differences scores approached, but did not reach, statistical significance (F= 3.68, p <.06). Table 8 reveals that on the immediate test the two groups had nearly identical scores, while the groups that were tested after two weeks had passed were significantly different.

Table 8: PV-CV difference scores by
time of test and level of violence

	High violence	Low violence	t
Immediate	−1.266	−1.233	0.12
Delayed	−1.800	−1.166	2.46 (p < .01)

The direction of the significant difference between the two delayed-test groups, however, is in the opposite direction of that predicted by H_2. These data indicate that the notion that violence might increase the relative recall of peripheral material had no support and may have been the reverse of the actual phenomenon.

In addition to the above findings, there were four significant interaction effects involving the time of testing. Three of these have already been presented, and the fourth will be discussed below. In general, time of test had an effect on recall; but the time at which the test was given was not a critical factor influencing the effects of violence level, color, and age group upon learning from a television program.

Other interactions

Scores on peripheral-visual learning produced several interaction effects. The interaction between format and violence level has been presented in Figure 3. In addition, there were interactions between format and age level (Figure 11) and format, age level, and violence level (Figure 12). The effects of format on peripheral-visual recall reversed themselves over the three age groups. Younger boys recalled more peripheral-visual material from a black and white presentation, while older boys recalled more from a color presentation. However, the three way interaction presented by Figure 12 indicates that the situation was not so straightforward. One group of fourth graders—B/W-Low Violence—seems to be completely out of line when compared to the other groups. The clearest finding seems to be that the color presentation of high aggressive violence (hollow triangles) produced consistently higher recall —over three grades—than the black and white version (solid triangles) of the same presentation.

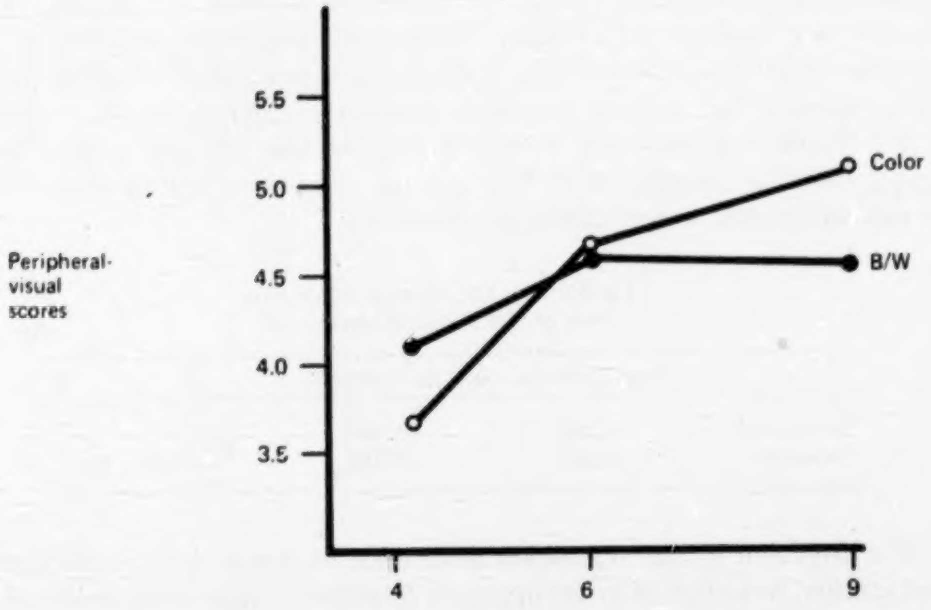


Figure 11: Interaction between format and grade on peripheral-visual scores

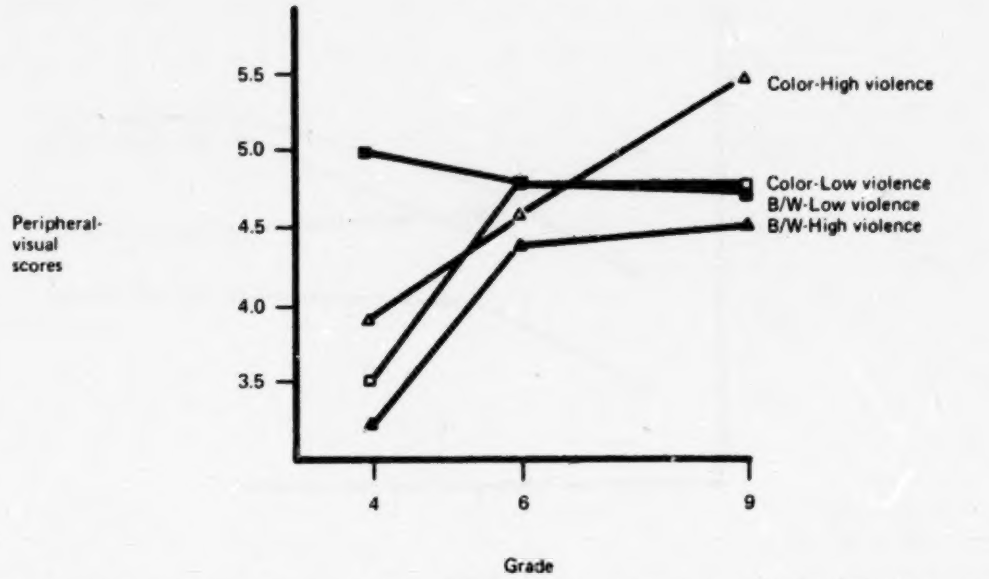


Figure 12: Interaction between format, violence level, and grade on peripheral-visual scores

The interactions among factors that produced different peripheral-visual scores may have been responsible for the significant interactions that were found for PV-CV difference scores. Figure 13 presents the interactions on this variable among format, violence level, and grade. Fourth graders displayed a completely mixed pattern: B/W-Low and Color-High presentations produced equivalently high peripheral-visual scores relative to central-visual scores. The Color-Low group was next highest, and the B/W-High group was lowest. Among the sixth graders the two predicted main effects seem to have occurred. Both Color groups are higher than both B/W groups, and both High Violence groups are higher than both Low Violence groups. The latter effect reverses in the ninth grade. The older boys scored higher when they saw the show in color; but they scored lower when they saw High Violence versions. Figure 14 presents interactions on PV-CV difference scores among format, violence level, and time of test. In this case one group (B/W-High-Delay) is very far from the other seven groups. It may have had an undue influence on the findings.

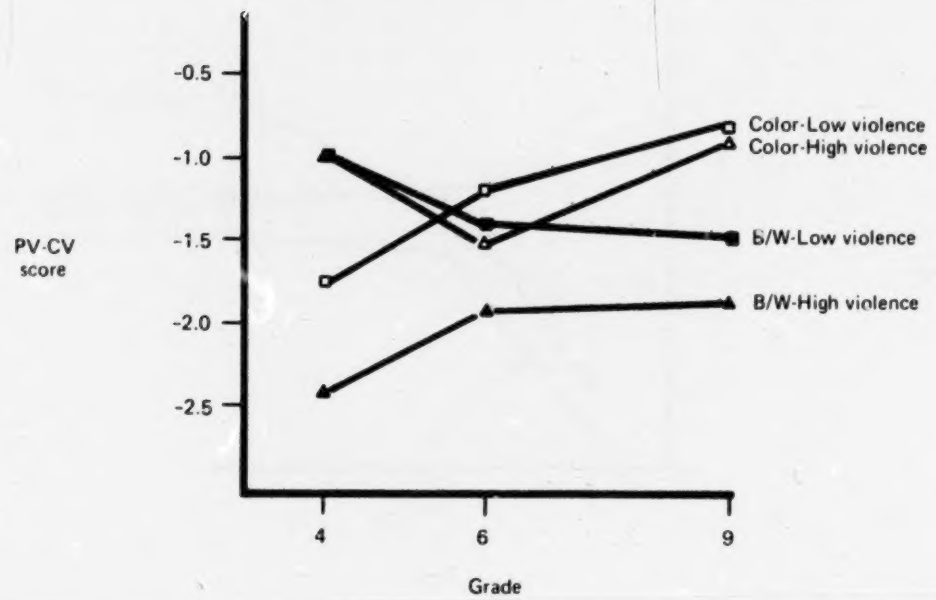


Figure 13: Interaction between format, violence level, and grade on PV-CV difference scores

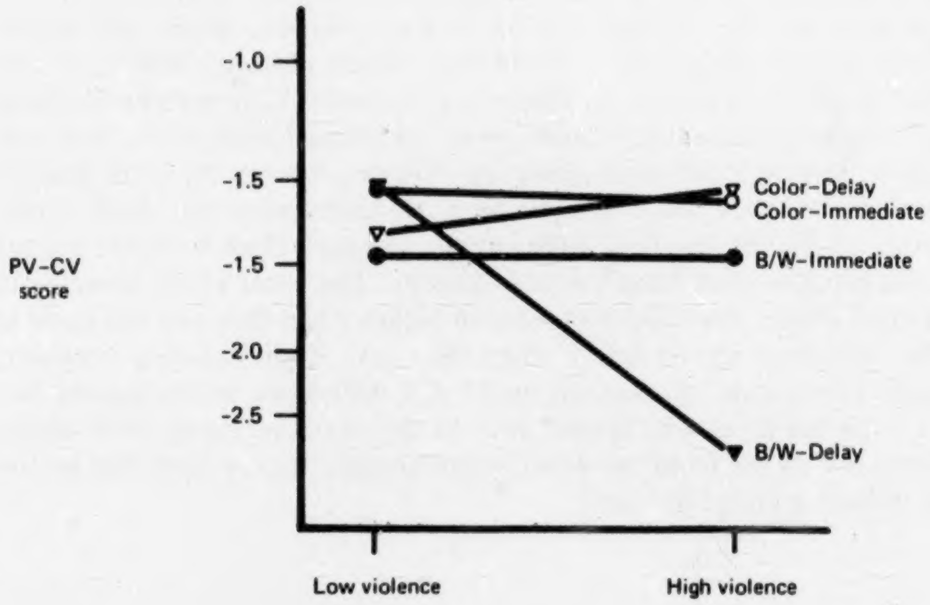


Figure 14: Interaction between format, violence level, and time of test on PV-CV difference scores

Figure 15 presents a significant four-way interaction effect on PA-CA difference scores. The only significant main effect on this variable was age level. From Figure 15 it is clear that PA-CA differences decreased significantly from the fourth to the sixth grade. Unfortunately, half of the ninth-grade groups displayed sharp rises above sixth-grade groups in the same experimental condition, and the other half displayed even sharper declines below the sixth-grade groups in the same experimental conditions. There is no meaningful pattern grouping the groups that rose and the groups that declined.

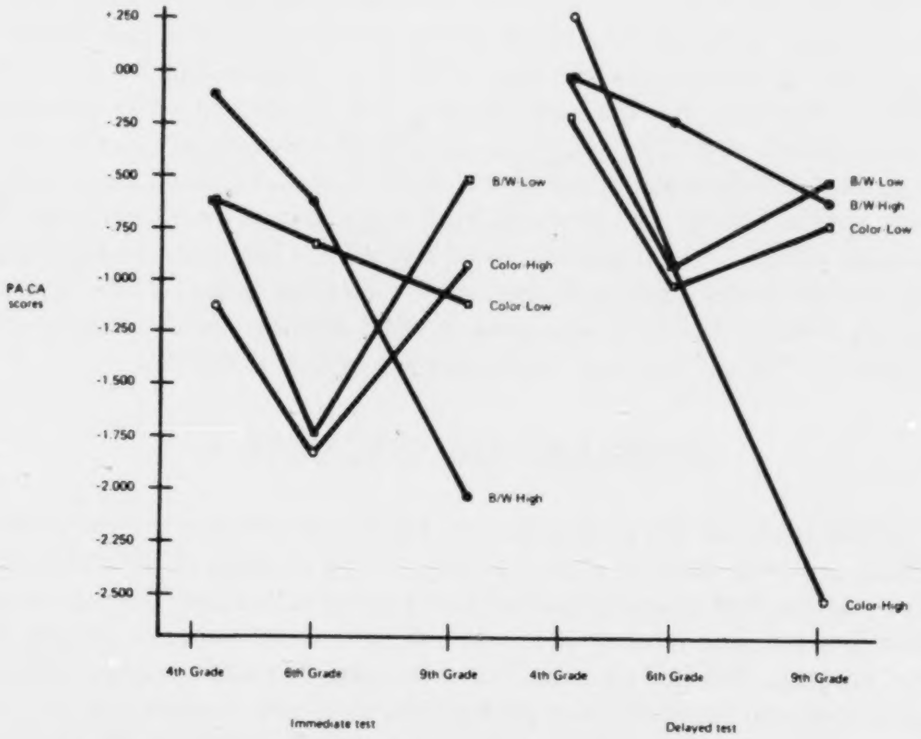


Figure 15: Four-way interaction effects on PA-CA difference scores

The special item

One recall item was included in the test; the correct answer indicated that nothing of the sort had happened. We included this item with the hope that wrong answers might indicate tendencies to fantasize about the program. The usual main effects for learning were present: older boys scored higher and those tested immediately scored higher. However, when the number of boys for each age group who had correctly answered the special item was compared with the number who would have been expected to get any item correct (total learning score divided by number of items multiplied by number of subjects per cell), the value of the resulting Chi-square test was far below statistical significance. The same failure to find differences occurred when violence levels and formats were compared. Results on the item simply corresponded to patterns of total recall.

The effects of home viewing habits

Two indicators of home viewing habits, amount viewed and presence of color television set, were included in the questionnaire. It was thought that the effects of color presentation in the experiment might be confounded with novelty effects for boys who did not have color television at home. Exactly 50 percent of the 240 subjects indicated that they had color television sets at home. (The A.C. Nielsen Station Index for 1970-71 indicates that the Lansing area had 51 percent color television sets in homes with television.) An analysis of covariance was calculated to control for presence of television in the home. Adjusted mean scores were within 1.5 percent of unadjusted mean scores on all variables for all main effects. A similar analysis of covariance indicated no significant effects due to the amount of television viewed at home. These findings are presumably due to an adequate random distribution of subjects with each type of home viewing pattern to each experimental cell.

SUMMARY AND DISCUSSION

The purpose of this project was to determine whether violent content affects learning from television among young viewers of different ages, and whether this learning is affected by color television. Boys from the fourth, sixth, and ninth grades were shown edited versions of a detective program that had appeared on television five years earlier. Results were analyzed for 80 boys in each grade. Half saw a version of the program with aggressive violence edited out; the other half saw a version of the program that included all the aggressive violence. The high and low violence versions were presented in color to half the boys, and in black and white to the others. One group in each grade-violence-color condition was tested for recall of program content immediately after the showing; another group was asked for opinions of the program immediately after the showing. Two weeks later, the group that had not been tested for recall was given a delayed recall test and the group that had not been tested for opinions was given a delayed opinion test.

Effects of violence

1. The High Violence presentation was perceived as significantly more violent than the Low Violence presentation.

2. Significantly more peripheral material was recalled from the Low Violence version.

3. There was no evidence that high aggressive violent content increased recall or raised the amount of peripheral material learned when central material learned is accounted for.

Findings 1, 2, and 3 were the results of pooling data over three age groups and two presentation formats. There were significant interaction effects.

Effects of color presentation

4. Color presentation produced a better recall rate for PV-CV difference scores than black and white presentation.

5. The boys who viewed the black and white presentations recalled more central-visual material than the boys who saw the color presentations.

6. The effect of shifting attention toward peripheral-visual material, relative to central visual material, (finding #4) was mainly due to higher PV-CV difference scores for color presentations of the High Violence program. Differences for the Low Violence presentation were small.

7. Color format had no significant overall effect on perceived violence in the programs.

8. Color format had no significant overall effect on how much the boys said they liked the program.

Differences across age groups

9. There were extremely significant differences between age groups on all categories of recall. The older boys recalled more than the younger boys.

10. Increases from fourth to sixth grade in recall were sharper than increases from sixth to ninth grade.

11. Fourth-grade boys recalled more peripheral material, relative to central material, than boys in the other two grades.

12. Older boys tended to focus more on central material than younger boys, for material presented by the sound track.

13. Younger boys were significantly more favorable in their evaluations of programs than older boys.

14. Fourth-grade boys liked the Low Violence version better than the High Violence version; older boys liked both versions equally.

15. For each level of violence, the younger boys said there was more violence in the programs than the older boys.

Interactions between violence level and grade

16. Younger boys recalled more from the Low Violence presentation, while older boys recalled more from the High Violence presentation. Sixth graders recalled about as much from each level of violence.

Differences between immediate and delayed tests

17. More was recalled immediately after presentation than two weeks later.

18. After a two-week period, the High Violence version was rated significantly more violent than it had been rated immediately after presentation. The Low Violence version was rated about the same at both times.

19. In the delayed test, the amount of peripheral-visual recalled relative to the amount of central visual material recalled was higher for the Low Violence presentation. The effect did not occur in the immediate test.

Implications

The effects of color format in this experiment were fairly consistent with earlier findings about the impact of color. Color does not improve the learning of the visual material that is central to a presentation. Thus, the cost involved in changing from a black and white instructional medium to a color system may not produce the expected benefits. (Of course, color is a necessary component for some subject matter.) On the other hand, the evidence indicated that in entertainment media the presence of color may alter the balance of central and peripheral material recalled. This phenomenon may have important side effects related to perceived consequences, rewards, and justification of aggressive violence.

Different age groups gave different responses to the programs. There was some support for the notion that older children are better able to 'focus' their attention on central aspects of a stimulus. More interesting was the finding that the fourth-grade boys liked the low violence version of the program better than the high violence version. At this age, violence is not an attractive thing. Older boys—perhaps more used to violence on television—liked both versions equally well. Younger boys also tended to say that a program was more violent. This may have been a verbal artifact of language use at different ages, but it also may have been a reflection of a greater sensitivity to violence among younger children.

Pooled over all age groups, the variable of aggressive violence had no important effects. (More peripheral material was recalled from the *low* violence version.) However, the violence-age interaction may point to important considerations for research on the effects of violence on children. The reversal of recall patterns (fourth grade: more from low violence version; sixth grade: no difference; ninth grade: more from high violence version) indicates that any measurements of supposed effects of violent content may be specific to the age group studied. It is possible that teenagers will be more likely to imitate violence (they recall more from violent programs) than younger children, who recall less and indicate less liking for violent presentations.

Limitations

This study was obviously limited in its generalizability. A single stimulus program was used; the effects of the variables studied may be different for different programs. A single demographic group was studied; the effects may not be the same for girls, less intelligent children, or children from lower socioeconomic backgrounds. These two features of the study would be crucial drawbacks if there were an attempt to create broad laws of behavior from the data. However, the study should be considered as a preliminary step. The areas under consideration have been given only minimal attention in the past. Perhaps these data can point toward further research regarding the effects of mass media on learning and behavior.

FOOTNOTES

1. The research upon which this report is based was performed pursuant to Contract No. HSM 42-70-68 with the National Institute of Mental Health, Health Services and Mental Health Administration, Department of Health, Education and Welfare.

Many people deserve credit for their work on this project. Jerry Buley was a capable and invaluable administrative assistant. Anita Immele made the computer lie down and roll over for us. Mr. Allan Silverbach, of 20th Century-Fox Television, was kind enough to let us use one of his programs for the project. The technicians at the National Educational Television videotape center in Ann Arbor provided critical assistance when it was time to edit the stimulus programs. Dr. Robert Docking, of the East Lansing public school system, encouraged us and allowed us to function smoothly and efficiently at all three grade levels. Support and cooperation were given by the following kind people: Diane Oldaugh, at East Lansing High School; Sal DeFranco, principal, and his assistant Craig Marsh, at McDonald Middle School; Brewster Lewis, principal of Bailey School, Sally Swartz, principal of Marble School; Charles Townsel, principal, and Lola O'Meara, of Pinecrest School. Donna Owen and Esther White were patient and competent experimenters throughout the project. Ed Amend, Gerry Hanneman, Jim Nyenhuis, and Ed Wotring helped to develop and refine the questionnaire. T. Harrell Allen, Akiba Cohen, Rogelio Cuyno, Phil Ericson, Peter Monge, Mike Rohla, Josep Rota, and Mark Steinberg all served capably in the field. Fred Henderson and Larry Cox, of Closed Circuit TV at M.S.U., managed to keep the color videotape equipment in tune, despite some repair jobs that had to be done on very short notice. Finally, thanks are in order to the secretaries and staff of the Department of Communication, Michigan State University; they have had to cope with rush orders and silly mistakes on our part.

The project could not have been done without the above people. I thank them all very much. But I have to admit that I am responsible, in the end, for this report. If you find oversights or blunders, they were mine.

—N.K.

REFERENCES

Baldwin, T. F., and Lewis, C. *Violence In Television: The Industry Looks at Itself,* Michigan State University (mimeo), April 1971.

Collins, W. A. Learning of media content: a developmental study. *Child Development,* 1970, **41**, 1133-42.

Deutschmann, P. J., Barrow, L. C., and McMillan, A. The efficiency of different modes of communication. *AV Communication Review,* 1961, **9**, 263-70.

Hale, G.A., Miller, L. K., and Stevenson, H. W. Incidental learning of film content: a developmental study. *Child Development,* 1968, **39**, 69-78.

Kanner, J. H., and Rosenstein, A. J. Television in army training: color vs. black and white. *AV Communication Review,* 1960, **8**, 243-52.

Katzman, N. I. *Aggressive violence: definitions and distinctions for mass media studies.* Michigan State University (mimeo), 1971 (submitted for publication).

Katzman, N. I., and Nyenhuis, J. Color vs. black and white effects on learning, opinion, and attention. 1971 (submitted for publication).

Kumata, H. Two studies in classroom teaching. In Schramm, W. (Ed.) *The impact of educational television.* Urbana: University of Illinois Press, 1960.

Larson, O. N., Grey, L. N., and Fortis, J. G. Goals and goal achievement methods in television content: models for anomie? *Sociological Inquiry,* 1963, **33**, 180-196.

Roberts, D. F. A developmental study of opinion change: source-orientation versus content orientation at three age levels. Unpublished doctoral dissertation, Stanford University, 1968.

Schaps, E., and Guest, L. Some pros and cons of color TV. *Journal of Advertising Research,* 1968, **8**, 28-39.

VanderMeer, A. W. Relative effectiveness of color and black and white in instructional films. Technical Report SDC 269-7-28. State College, Pa.: Instructional Film Research Program, The Pennsylvania State University, 1952.

Appendix A: Summaries of low violence and high violence versions of the program used in the project (Summaries are presented in parallel columns for comparison.)

Low Violence—20 minutes, 16 seconds	*High Violence*—20 minutes, 25 seconds
A United Air Lines jet lands at a large metropolitan airport. A young man walks through the airport to meet a stewardess who runs to meet him as she comes from the plane. They kiss. He jokingly comments that it seems like years since he has seen her. She replies that it has been exactly 53 hours and 30 minutes.	A United Air Lines jet lands at a large metropolitan airport. A young man walks through the airport to meet a stewardess who runs to meet him as she comes from the plane. They kiss. He jokingly comments that it seems like years since he has seen her. She replies that it has been exactly 53 hours and 30 minutes.
In the hallway by her apartment door (number 9) the couple is happily talking. He asks about her trip, and she replies that after fetching and carrying for 6,000 miles a house and a husband would be a cinch to take care of. He asks, "Any prospects?" They kiss again.	In the hallway by her apartment door (number 9) the couple is happily talking. He asks about her trip, and she replies that after fetching and carrying for 6,000 miles a house and a husband would be a cinch to take care of. He asks, "Any prospects?" They kiss again.
Loud and ominous music. We see a man hurriedly searching through things in a bedroom. He hears a noise in the next room, stops, and draws a gun.	Loud and ominous music. We see a man hurriedly searching through things in a bedroom. He hears a noise in the next room, stops, and draws a gun.
The couple from the airport enter the living room. He proposes dinner at a romantic restaurant. She replies, "Oh, Jim, could we . . ?" He says that first he has to check in because he is on standby duty. He goes to the telephone while she goes to the bedroom to change.	The couple from the airport enter the living room. He proposes dinner at a romantic restaurant. She replies, "Oh, Jim, could we. . .?" He says that first he has to check in because he is on standby duty. He goes to the telephone while she goes to the bedroom to change.
	From inside the bedroom there is the shadow of a gun on the door. Ominous music. The girl enters and the man inside shuts the door, grabs her violently, and holds his hand over her mouth. He says, "Get rid of that guy fast! I mean it, baby."

Low Violence—Continued

In a police station a uniformed policeman answers the telephone. Jim, in the apartment, says, "Dad, it's me." He wants to know if he has to work tonight. His dad asks Sam to talk to Jim. Sam tells Jim that, "We're still the backup team." Jim is displeased, and says that he will be at Cloris's place if he is needed.

Cloris comes out of the bedroom as Jim is putting down the phone. He says that he's still on call, and suggests that they have supper in the apartment. Cloris tensely replies that she thinks they should forget dinner tonight. She is tired. Jim says that what she needs is a plate of spaghetti. Cloris says she is not hungry and that he should go. Jim is upset. Cloris asks him to call tomorrow. He agrees and leaves unhappily.

Cloris turns to the bedroom door. It opens and the man comes out. Ominous music. He walks up behind her.

Theme music comes on. Program titles and the names of the actors are superimposed over scenes of a city at night and a police car cruising through the streets.

Inside the police station Jim and his father are standing by a desk as Sam comes up and tells Jim that they have the rest of the night off. Dad tells Jim to go make up with his girl. Jim asks Sam to drop him off at Cloris's apartment.

In the apartment Cloris and the man are arguing. She has not heard from him in four years. He goes to her, but she gets up and moves away. She says, "We're divorced, Wade." In the dialogue we learn that she has gotten a Mexican divorce and that her lawyers have served Wade with the papers. He denies having received them; but she insists that they are divorced and she has a new life to lead. She is going to marry Jim the minute he asks her. Wade discovers that Cloris has not told Jim about the first marriage and threatens to

High Violence—Continued

Cloris comes out of the bedroom as Jim is putting down the phone. He says that he's still on call, and suggests that they have supper in the apartment. Cloris tensely replies that she thinks they should forget dinner tonight. She is tired. Jim says that what she needs is a plate of spaghetti. Cloris says she is not hungry and that he should go. Jim is upset. Cloris asks him to call tomorrow. He agrees and leaves unhappily.

Cloris turns to the bedroom door. It opens and the man comes out. Ominous music. He walks up behind her.

The man grabs Cloris by the shoulders. She struggles, but he forces a kiss. She pushes him away and slaps him in the face. He yells, "What's the matter? Is that any way to welcome your husband?"

Theme music comes on. Program titles and the names of the actors are superimposed over scenes of a city at night and a police car cruising through the streets.

Inside the police station Jim and his father are standing by a desk as Sam comes up and tells Jim that they have the rest of the night off. Dad tells Jim to go make up with his girl. Jim asks Sam to drop him off at Cloris's apartment.

In the apartment Cloris and the man are arguing. She has not heard from him in four years. He goes to her, but she gets up and moves away. She says, "We're divorced, Wade." In the dialogue we learn that she has gotten a Mexican divorce and that her lawyers have served Wade with the papers. He denies having received them; but she insists that they are divorced and she has a new life to lead. She is going to marry Jim the minute he asks her. Wade discovers that Cloris has not told Jim about the first marriage and threatens to

Low Violence—Continued

spoil her plans with Jim if she refuses to help. He grips her by the shoulders.

The buzzer from downstairs sounds. Cloris breaks away and pushes an intercom button. Downstairs a man in a suit is biting his nails while waiting at the intercom. He says, "Tell Wade I know he's up there, tell him not to hurry. I'll be around..............Like always."

In the apartment, Wade looks worried as he hears the man. He goes to the window and sees the man looking up. The man steps into the passenger side of a car and waits there. Cloris guesses that Wade is in trouble. She asks who the man is. Wade replies that he is a killer. Wade had pulled a warehouse job with him and his brother. He picks up a glass and takes a drink of something alcoholic.

"Well...now you admit you're a thief," says Cloris. Wade yells that she must have known that when they were married. Cloris, emphatically replies that she was stupid to believe all of his lies.

Wade explains that the kid brother had been killed during the warehouse job and the older brother swore to kill Wade. Wade has been running from him, but this morning he saw the man checking into a hotel across the street from Wade's. Cloris suggests that Wade go to the police. He says, "Not me, baby; somehow I always luck out."

In the street outside the apartment building the man is sitting in his car as Sam and Jim pull up. They have an unmarked car, but the police radio is broadcasting a message. Hearing this, the man moves to the driver's seat and drives away.

Jim invites Sam to come up and meet Cloris. Sam agrees, "but just for a minute."

Upstairs, Cloris denies that she has money. But Wade has found a bankbook and shouts that she has $1,700. He wants all of it. He grabs her arm, and just then the buzzer sounds in a unique code.

High Violence—Continued

spoil her plans with Jim if she refuses to help. He grips her by the shoulders.

The buzzer from downstairs sounds. Cloris breaks away and pushes an intercom button. Downstairs a man in a suit is biting his nails while waiting at the intercom. He says, "Tell Wade I know he's up there, tell him not to hurry. I'll be around..............Like always."

In the apartment, Wade looks worried as he hears the man. He goes to the window and sees the man looking up. The man steps into the passenger side of a car and waits there. Cloris guesses that Wade is in trouble. She asks who the man is. Wade replies that he is a killer. Wade had pulled a warehouse job with him and his brother. He picks up a glass and takes a drink of something alcoholic.

"Well...now you admit you're a thief," says Cloris. Wade yells that she must have known that when they were married. Cloris, emphatically replies that she was stupid to believe all of his lies.

Wade yells and throws the glass with the drink across the room, where it smashes noisily.

Wade explains that the kid brother had been killed during the warehouse job and the older brother swore to kill Wade. Wade has been running from him, but this morning he saw the man checking into a hotel across the street from Wade's. Cloris suggests that Wade go to the police. He says, "Not me, baby; somehow I always luck out."

In the street outside the apartment building the man is sitting in his car as Sam and Jim pull up. They have an unmarked car, but the police radio is broadcasting a message. Hearing this, the man moves to the driver's seat and drives away.

Jim invites Sam to come up and meet Cloris. Sam agrees, "but just for a minute."

Upstairs, Cloris denies that she has money. But Wade has found a bankbook and shouts that she has $1,700. He wants all of it. He grabs her arm, and just then the buzzer sounds in a unique code.

Low Violence—Continued	*High Violence*—Continued
Cloris says, "It's Jim." Wade lets her go and suggests that she introduce Jim to her "loving husband."	Cloris says, "It's Jim." Wade lets her go and suggests that she introduce Jim to her "loving husband."
Cloris opens the door, on which apartment number 9 is again visible. Jim and Sam are there. Jim sees Wade and is surprised. He introduces Sam. Cloris hesitantly says, "This is...Wade."	Cloris opens the door, on which apartment number 9 is again visible. Jim and Sam are there. Jim sees Wade and is surprised. He introduces Sam. Cloris hesitantly says, "This is...Wade."
Wade says Cloris has told him a lot about Jim; but Jim responds, "she hasn't told me about you." Wade is being suave and friendly. "What's the matter Cloris? You ashamed of your brother?" Cloris introduces Wade as her brother. Everyone is introduced to everyone. Full names are Jim Briggs, Sam Stone, Wade Harte and Cloris Harte.	Wade says Cloris has told him a lot about Jim; but Jim responds, "she hasn't told me about you." Wade is being suave and friendly. "What's the matter Cloris? You ashamed of your brother?" Cloris introduces Wade as her brother. Everyone is introduced to everyone. Full names are Jim Briggs, Sam Stone, Wade Harte and Cloris Harte.
Wade asks if Sam and Jim are in business together. Sam replies that they are both police officers. Wade, holding back his surprise, says that police are mighty important these days and suggests a drink. Sam excuses himself and leaves. Wade and Jim go into the apartment to get acquainted.	Wade asks if Sam and Jim are in business together. Sam replies that they are both police officers. Wade, holding back his surprise, says that police are mighty important these days and suggests a drink. Sam excuses himself and leaves. Wade and Jim go into the apartment to get acquainted.
Wade claims that he is a traveling salesman, selling anything and everything. Then he suggests that Cloris make them all dinner. Jim has a better idea; he will cook spaghetti. Jim takes off his sport jacket, puts his gun and badge in the jacket pocket, and hangs it in the closet. Then he invites Cloris into the kitchen "to help the chef."	Wade claims that he is a traveling salesman, selling anything and everything. Then he suggests that Cloris make them all dinner. Jim has a better idea; he will cook spaghetti. Jim takes off his sport jacket, puts his gun and badge in the jacket pocket, and hangs it in the closet. Then he invites Cloris into the kitchen "to help the chef."
When they leave, Wade's face turns serious. He checks the window for the man in the car. He gets up and moves toward the closet, calling to the kitchen "Hey, I'm going to run out and get us a good bottle of wine. Need anything?" Wade opens the closet and removes the gun and badge from Jim's jacket. Jim yells, "No thanks," from the kitchen. Wade leaves.	When they leave, Wade's face turns serious. He checks the window for the man in the car. He gets up and moves toward the closet, calling to the kitchen "Hey, I'm going to run out and get us a good bottle of wine. Need anything?" Wade opens the closet and removes the gun and badge from Jim's jacket. Jim yells, "No thanks," from the kitchen. Wade leaves.
The next scene is in the lobby of a cheap hotel. Wade flashes the police badge at the desk clerk and demands to know the room number of Paulie Egan. After convincing the clerk that he does not intend to make an arrest he gets the information and goes to the stairs.	The next scene is in the lobby of a cheap hotel. Wade flashes the police badge at the desk clerk and demands to know the room number of Paulie Egan. After convincing the clerk that he does not intend to make an arrest he gets the information and goes to the stairs.

Low Violence—Continued	*High Violence*—Continued
The desk clerk phones Egan's room and tells Egan that there's a cop on the way up to see him. Egan is the man who had been after Wade. He hides a revolver in a wastebasket.	
In Egan's hotel room the radio is playing loudly and Egan is shaving with a safety razor in the bathroom. Wade opens the door quietly with a gun in one hand. He picks up a pillow and knocks on the door from the inside. Egan calls, "It's open..." from the bathroom. Wade turns up the volume on the radio, holds the pillow in front of the gun to muffle the sound of the shots, and prepares to fire. Egan turns from the shaving mirror, sees Wade, and displays a look of shock.	In Egan's hotel room the radio is playing loudly and Egan is shaving with a safety razor in the bathroom. Wade opens the door quietly with a gun in one hand. He picks up a pillow and knocks on the door from the inside. Egan calls, "It's open..." from the bathroom. Wade turns up the volume on the radio, holds the pillow in front of the gun to muffle the sound of the shots, and prepares to fire. Egan turns from the shaving mirror, sees Wade, and displays a look of shock.
	Wade fires the gun through the pillow. Egan is hit. He staggers from the bathroom, knocking over some shaving articles. Wade fires another into his chest at point blank range. The bullet holes can be seen against Egan's tee shirt. Egan staggers backward and collapses to the floor. Wade kicks him onto his back, again the bullet holes and blood can be seen on Egan's chest. Then Wade steps out the window on to a fire escape.
First there is a scene of cars passing by the front of the hotel. It is day time. Then a flashbulb fires, and there are policemen carrying a stretcher with a corpse past Jim. The desk clerk is telling Sam that, "it was one of your men who killed Paulie Egan. . .a plainclothes cop...he flashed his I.D. and I make it a point never to forget a badge number. His was 1348." Sam registers shock and confronts Jim. He wants to know what happened after he left Jim at Cloris's last night. Jim says nothing happened. Sam says that the man who murdered Egan used Jim's badge and maybe even Jim's gun. He says Egan was shot with a .38. Jim says he had checked his gun and it was fully loaded. He smells the chambers. He tells Sam that it *has* been fired recently, reloaded and replaced in his pocket. There was only one man who could have done it.	First there is a scene of cars passing by the front of the hotel. It is daytime. Then a flashbulb fires, and there are policemen carrying a stretcher with a corpse past Jim. The desk clerk is telling Sam that, "it was one of your men who killed Paulie Egan...a plainclothes cop...he flashed his I.D. and I make it a point never to forget a badge number. His was 1348." Sam registers shock and confronts Jim. He wants to know what happened after he left Jim at Cloris's last night. Jim says nothing happened. Sam says that the man who murdered Egan used Jim's badge and maybe even Jim's gun. He says Egan was shot with a .38. Jim says he had checked his gun and it was fully loaded. He smells the chambers. He tells Sam that it *has* been fired recently, reloaded and replaced in his pocket. There was only one man who could have done it.

Low Violence—Continued

Sam and Jim burst through the door to Cloris's apartment. Cloris is surprised. The two men look around quickly and say nothing until they have searched the apartment. Sam angrily demands where her brother is, but Jim is cooler and says he will talk to her. Sam goes to phone in an "APB" while Jim explains to Cloris that her "brother" used the badge and gun to murder a man. She says she does not know where Wade is, and Jim accuses her of lying. She says she had no idea of what he was up to and that he left last night right after Jim. Jim angrily replies, "You had no idea...he's your brother!"

Cloris very dramatically admits, "He's not my brother...he's my husband...my ex-husband." Jim sadly says, "You lied...Why?" "Don't you know?" she asks. The scene fades out.

Wade telephones Cloris from a booth. Cloris answers her phone after it rings. Jim walks up to Cloris and she holds the phone so that they both can hear what Wade says. Wade asks if she has the money. She says not yet. He says the bank will open in 20 minutes and she had better get the money then. He says she had better bring the whole kitty. Jim covers the mouthpiece and whispers, "Where?" Cloris asks, and Wade tells her to walk across the plaza and through the tunnel after she leaves the bank. He hangs up.

Jim: "I can't ask you to play along with this."

Cloris: "Just tell me what you want me to do."

Tightly clutching her handbag, Cloris leaves the bank and begins to walk across a shopping plaza mall. Jim and Sam are waiting in an unmarked car.

Jim's dad is standing near the bank, wearing what seems to be a hearing aid. Dad puts a cigarette to his mouth and begins speaking into a microphone in his palm. "Special unit one this is 'Stroller' ...come in please." Sam replies on the police radio, and Dad describes Cloris as she walks across the plaza. He tells Sam and Jim to stay out of sight until they are called, Wade knows them.

Cloris nervously and slowly walks across the plaza. "Stroller"

High Violence—Continued

Sam and Jim burst through the door to Cloris's apartment. Cloris is surprised. The two men look around quickly and say nothing until they have searched the apartment. Sam angrily demands where her brother is, but Jim is cooler and says he will talk to her. Sam goes to phone in an "APB" while Jim explains to Cloris that her "brother" used the badge and gun to murder a man. She says she does not know where Wade is, and Jim accuses her of lying. She says she had no idea of what he was up to and that he left last night right after Jim. Jim angrily replies, "You had no idea...he's your brother!"

Cloris very dramatically admits, "He's not my brother...he's my husband...my ex-husband." Jim sadly says, "You lied...Why?" "Don't you know?" she asks. The scene fades out.

Wade telephones Cloris from a booth. Cloris answers her phone after it rings. Jim walks up to Cloris and she holds the phone so that they both can hear what Wade says. Wade asks if she has the money. She says not yet. He says the bank will open in 20 minutes and she had better get the money then. He says she had better bring the whole kitty. Jim covers the mouthpiece and whispers, "Where?" Cloris asks, and Wade tells her to walk across the plaza and through the tunnel after she leaves the bank. He hangs up.

Jim: "I can't ask you to play along with this."

Cloris: "Just tell me what you want me to do."

Jim's dad is standing near the bank, wearing what seems to be a hearing aid. Dad puts a cigarette to his mouth and begins speaking into a microphone in his palm. "Special unit one this is 'Stroller'...come in please." Sam replies on the police radio, and Dad describes Cloris as she walks across the plaza. He tells Sam and Jim to stay out of sight until they are called, Wade knows them.

Low Violence—Continued	*High Violence*—Continued
walks behind her, pretending that he is on his way somewhere. Cloris keeps looking around for something.	
Cloris reaches the steps to an underground tunnel for pedestrians and begins to go down the steps. Wade is across the street; he sees her and runs to the far end of the tunnel. Cloris looks back at "Stroller" for assurance. "Stroller" reports to Sam and Jim that "she's coming to the mall now." As Cloris approaches the end of the tunnel a voice calls her name. "Stroller" reports the contact, and Sam says that they will move in from above.	Cloris reaches the steps to an underground tunnel for pedestrians and begins to go down the steps. Wade is across the street; he sees her and runs to the far end of the tunnel. Cloris looks back at "Stroller" for assurance. "Stroller" reports to Sam and Jim that "she's coming to the mall now." As Cloris approaches the end of the tunnel a voice calls her name. "Stroller" reports the contact, and Sam says that they will move in from above.
At the mouth of the tunnel Wade takes Cloris by the arm and moves away up a flight of stairs. Sam and Jim pull their car to a screeching halt and jump out. They pull their guns and begin to run in two directions to surround Wade. Wade has not seen them yet. Cloris nervously shoves the money at him, but he will not let her go. Sam runs toward Wade with his gun drawn and shouts "Hold it! Police!"	At the mouth of the tunnel Wade takes Cloris by the arm and moves away up a flight of stairs. Sam and Jim pull their car to a screeching halt and jump out. They pull their guns and begin to run in two directions to surround Wade. Wade has not seen them yet. Cloris nervously shoves the money at him, but he will not let her go. Sam runs toward Wade with his gun drawn and shouts "Hold it!" Police!"
	Wade pulls Cloris in front of him as a shield, draws his gun, and fires at Sam. Bystanders are frantically screaming and trying to get out of the way. Wade runs away from Sam, but he sees Jim coming from the other direction. Jim pushes people out of the way as Wade fires at Jim.
Wade jumps over a railing to a lower level of the plaza. Jim and Sam pursue him from two different sets of stairs. Wade is using people and objects as shields, so the police cannot get a shot at him.	Wade jumps over a railing to a lower level of the plaza. Jim and Sam pursue him from two different sets of stairs. Wade is using people and objects as shields, so the police cannot get a shot at him.
	Wade fires at Sam, and Jim fires at Wade from the side. Wade returns Jim's fire.
Wade runs into a large fountain in the center of the mall and tries to take cover amid a large modern sculpture. Jim chases after him.	Wade runs into a large fountain in the center of the mall and tries to take cover amid a large modern sculpture. Jim chases after him.
	Wade fires at Jim from the fountain. Jim takes cover and shoots back. "Stroller" and Sam converge on the fountain. Wade shoots at Sam. Sam and Stroller shoot at Wade. Wade tries to fire again, but his gun does not work.
Still in the fountain by the centerpiece, Wade is urgently working	Still in the fountain by the centerpiece, Wade is urgently working

Low Violence—Continued	*High Violence*—Continued
at his gun. Jim comes up behind him and knocks the gun from his hand with a karate chop.	at his gun. Jim comes up behind him and knocks the gun from his hand with a karate chop.
	Jim and Wade have a violent fistfight in the water. Wade knocks Jim down and tries to run, but Jim catches him and delivers the final blows.
Sam and "Stroller" come running to the edge of the fountain as Jim is dragging Wade to them. Cloris is there. Jim, dripping wet, looks at her. She looks dramatically at him, turns, and walks away. There is a closeup of Jim. The music comes up and out.	Sam and "Stroller" come running to the edge of the fountain as Jim is dragging Wade to them. Cloris is there. Jim, dripping wet, looks at her. She looks dramatically at him, turns, and walks away. There is a closeup of Jim. The music comes up and out.

Appendix B: Questions used in recall test

Central - visual items (x indicates correct alternative)

When Jim gets ready to cook, he puts his gun in. . .
( ) HIS SHOULDER HOLSTER
( ) HIS BELT
(x) HIS COAT
( ) THE BEDROOM

Before leaving the apartment, the killer takes Jim's . . .
( ) MONEY AND IDENTIFICATION
( ) GUN AND MONEY
(x) IDENTIFICATION AND GUN
( ) GUN ONLY

Wade got information from the hotel desk clerk by . . .
(x) SHOWING A BADGE
( ) SHOWING HIS GUN
( ) THREATENING TO SHOOT THE CLERK
( ) THREATENING TO ARREST THE CLERK

Egan was shot by . . .
( ) DAN
( ) JIM
( ) SAM
(x) WADE

Sam and Jim break into the apartment. They look around, and . . .
( ) THEY CAN'T FIND CLORIS
( ) THEY FIND CLUES
(x) THEY CAN'T FIND WADE
( ) THEY FIND THE MONEY

The police know where Wade wants Cloris to bring the money. This is because Jim . . .
( ) LISTENS ON ANOTHER TELEPHONE
( ) GUESSES THE RIGHT PLACE
( ) OVERHEARS THEM FROM THE KITCHEN
(x) SHARES A TELEPHONE WITH CLORIS

At the end of the show, the killer is . . .
( ) DROWNED
( ) SLAPPED BY CLORIS
(x) CAPTURED
( ) WOUNDED

Central-auditory items (x indicates correct alternative)

The man who follows Wade to Cloris's apartment had told Wade. . .
( ) HE WANTED MONEY
(x) HE WAS GOING TO KILL WADE

( ) HE WOULD MEET WADE AT THE HOTEL
( ) HE WAS WADE'S BROTHER

Egan wants to see Wade because . . .
( ) THEY ARE PLANNING A ROBBERY
(x) EGAN'S BROTHER WAS KILLED
( ) EGAN JUST GOT OUT OF JAIL
( ) WADE CHEATED HIM

Wade wants Cloris to . . .
( ) GET A DIVORCE
( ) HIDE HIM IN HER APARTMENT
( ) TRY TO ESCAPE WITH HIM
(x) GIVE HIM MONEY

Wade tries to get Cloris to help him by threatening to . . .
( ) SHOOT HER
( ) TAKE HER MONEY
(x) SPOIL HER PLANS WITH JIM
( ) REFUSE TO DIVORCE HER

Wade introduces himself to Jim and Sam as Cloris's . . .
(x) BROTHER
( ) HUSBAND
( ) BOYFRIEND
( ) EX-HUSBAND

The night clerk identifies the man who killed Egan from the . . .
(x) BADGE NUMBER
( ) DESCRIPTION SAM GIVES HIM
( ) KILLER'S NAME
( ) PHOTOGRAPH HE SEES

After the killer gives her instructions, Cloris tells Jim that she is . . .
( ) AFRAID
(x) WILLING TO DO WHAT HE WANTS
( ) NERVOUS
( ) IN LOVE WITH HIM

When Cloris realizes what Wade has done, she . . .
(x) AGREES TO HELP THE POLICE
( ) PHONES FOR JIM
( ) TELLS WADE TO GO AWAY
( ) BEGINS TO CRY

Peripheral-visual items (x indicates correct alternative)

The airline Cloris arrives on is . . .
(x) UNITED
( ) AMERICAN
( ) PAN AMERICAN
( ) TWA

Cloris works as a . . .
( ) REPORTER
( ) MODEL
(x) STEWARDESS
( ) POLICE WOMAN

When Egan is sitting in his car, he is wearing a . . .
( ) COAT
( ) SWEATER
( ) SHIRT AND TIE
(x) SUIT

Cloris's apartment number is . . .
( ) 5
( ) 7
(x) 9
( ) 11

When we see the flash from a camera . . .
( ) SAM AND JIM ARE TALKING TO A REPORTER
( ) SOMEONE IS TAKING A PICTURE OF CLORIS
( ) WADE IS PRETENDING TO BE A POLICEMAN
(x) THEY ARE CARRYING OUT EGAN'S BODY

Cloris carries the money from the bank in a . . .
(x) PURSE
( ) PAPER BAG
( ) BRIEFCASE
( ) LARGE ENVELOPE

Sam and Jim's car looks like a . . .
( ) POLICE CAR
( ) STATION WAGON
( ) CONVERTIBLE
(x) NORMAL CAR

Peripheral-auditory items (x indicates correct alternative)

Cloris and Jim have been apart for . . .
(x) A COUPLE OF DAYS
( ) A HALF HOUR
( ) TWO MONTHS
( ) A YEAR

In her bank account, Cloris has . . .
( ) $500
( ) $700
( ) $1000
(x) $1700

Cloris and Wade argue in her apartment when Jim comes to the door. Before she answers the door, Cloris knows it is Jim because . . .

( ) SHE EXPECTED HIM
(x) HE BUZZED FROM DOWNSTAIRS
( ) HE CALLED UP TO THE WINDOW
( ) HE TELEPHONED HER

Wade says his job is . . .

( ) DETECTIVE
( ) INSURANCE MAN
(x) TRAVELING SALESMAN
( ) BUSINESSMAN

The kind of food Jim makes for dinner is . . .

( ) STEAK
( ) HAMBURGER
( ) PORK AND BEANS
(x) SPAGHETTI

While Jim and Cloris are cooking supper, Wade leaves. He says he is going to get some . . .

( ) FRESH AIR
( ) CIGARETTES
(x) WINE
( ) ICE CREAM

As Cloris goes to meet the killer, a police agent follows her. His code name is . . .

( ) ROVER
( ) WALKER
( ) STRANGER
(x) STROLLER

The police car arrives, Sam jumps out and yells at the killer . . .

( ) "GET YOUR HANDS UP!"
(x) "HOLD IT—POLICE!"
( ) "ALL RIGHT, DROP IT!"
( ) "STOP RIGHT THERE!"

Appendix C: Tables of complete four-way analysis of variance for major dependent variables

Table C-1: Index of perceived violence

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	3.504	1	3.504	1.60	
Violence level	33.004	1	33.004	15.07	.0005
Grade	21.775	2	10.887	4.97	.008
Time of test	10.837	1	10.837	4.95	.027
Color x Violence	.937	1	.937	*	
Color x Grade	5.608	2	2.804	1.28	
Color x Time	3.504	1	3.504	1.60	
Violence x Grade	1.058	2	.529	*	
Violence x Time	3.037	1	3.037	1.38	
Grade x Time	5.275	2	2.637	1.20	
Color x Violence x Grade	9.525	2	4.762	2.17	
Color x Violence x Time	3.504	1	3.504	1.60	
Color x Grade x Time	1.108	2	.554	*	
Violence x Grade x Time	1.225	2	.612	*	
Color x Violence x Grade x Time	1.358	2	.679	*	
Within groups	472.899	216	2.189		
Total	578.162	239			

*Value of F less than 1.00

Table C-2: Index of liking of program

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	4.004	1	4.004	*	
Violence level	3.037	1	3.037	*	
Grade	216.508	2	108.254	20.14	.0005
Time of test	3.504	1	3.504	*	
Color x Violence	4.537	1	4.537	*	
Color x Grade	4.008	2	2.004	*	
Color x Time	3.037	1	3.037	*	
Violence x Grade	14.175	2	7.087	1.32	
Violence x Time	1.504	1	1.504	*	
Grade x Time	11.558	2	5.779	1.08	
Color x Violence x Grade	4.075	2	2.037	*	
Color x Violence x Time	4.004	1	4.004	*	
Color x Grade x Time	2.025	2	1.012	*	
Violence x Grade x Time	29.258	2	14.629	2.72	
Color x Violence x Grade x Time	5.658	2	2.829	*	
Within groups	1161.100	216			
Total	1471.995	239			

*Value of F less than 1.00

Table C-3: Total recall among all items

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	6.666	1	6.666	*	
Violence level	14.016	1	14.016	1.11	
Grade	987.608	2	493.804	39.25	.0005
Time of test	952.016	1	952.016	75.67	.0005
Color x Violence	26.666	1	26.666	2.11	
Color x Grade	70.658	2	35.329	2.80	
Color x Time	6.666	1	6.666	*	
Violence x Grade	99.258	2	49.629	3.94	.021
Violence x Time	2.016	1	2.016	*	
Grade x Time	11.608	2	5.804	*	
Color x Violence x Grade	6.608	2	3.304	*	
Color x Violence x Time	8.066	1	8.066	*	
Color x Grade x Time	9.558	2	4.779	*	
Violence x Grade x Time	56.058	2	28.029	2.23	
Color x Violence x Grade x Time	8.108	2	4.054	*	
Within groups	2717.600	216	12.581		
Total	4983.183	239			

*Value of F less than 1.00

Table C-4: Recall of all peripheral items

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	.266	1	.266	*	
Violence level	14.016	1	14.016	3.93	.05
Grade	185.275	2	92.637	25.97	.0005
Time of test	220.416	1	220.416	61.78	.0005
Color x Violence	10.416	1	10.416	2.91	
Color x Grade	20.908	2	10.454	2.93	
Color x Time	2.016	1	2.016	*	
Violence x Grade	14.558	2	7.279	2.04	
Violence x Time	.066	1	.066	*	
Grade x Time	6.558	2	3.279	*	
Color x Violence x Grade	14.058	2	7.029	1.97	
Color x Violence x Time	.066	1	.066	*	
Color x Grade x Time	.358	2	.179	*	
Violence x Grade x Time	33.408	2	16.704	4.68	.01
Color x Violence x Grade x Time	1.608	2	.804	*	
Within groups	770.600	216	3.567		
Total	1294.600	239			

*Value of F less than 1.00

Table C-5: Recall of peripheral auditory items

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	.937	1	.937	*	
Violence level	3.037	1	3.037	2.10	
Grade	54.658	2	27.329	18.93	.0005
Time of test	34.504	1	34.504	23.89	.0005
Color x Violence	.704	1	.704	*	
Color x Grade	2.725	2	1.362	*	
Color x Time	1.204	1	1.204	*	
Violence x Grade	2.275	2	1.137	*	
Violence x Time	2.204	1	2.204	1.53	
Grade x Time	.758	2	.379	*	
Color x Violence x Grade	.308	2	.154	*	
Color x Violence x Time	3.037	1	3.037	2.10	
Color x Grade x Time	3.158	2	1.579	1.09	
Violence x Grade x Time	13.808	2	6.904	4.78	.01
Color x Violence x Grade x Time	6.775	2	3.387	2.35	
Within groups	311.900	216	1.443		
Total	441.995	239			

*Value of F less than 1.00

Table C-6: Recall of peripheral visual items

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	.204	1	.204	*	
Violence level	4.004	1	4.004	2.88	
Grade	39.433	2	19.716	14.19	.0005
Time of test	80.504	1	80.504	57.94	.0005
Color x Violence	16.537	1	16.537	11.90	.001
Color x Grade	9.033	2	4.516	3.25	.04
Color x Time	.104	1	.104	*	
Violence x Grade	8.233	2	4.116	2.96	
Violence x Time	3.037	1	3.037	2.18	
Grade x Time	3.433	2	1.716	1.23	
Color x Violence xGrade	10.300	2	5.150	3.71	.026
Color x Violence x Time	4.004	1	4.004	2.88	
Color x Grade x Time	2.133	2	1.066	*	
Violence x Grade x Time	4.300	2	2.150	1.55	
Color x Violence x Grade x Time	2.533	2	1.266	*	
Within groups	300.099	216	1.389		
Total	487.895	239			

*Value of F less than 1.00

Table C-7: Recall of all central items

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	4.266	1	4.266	*	
Violence level	.000	1	.000	*	
Grade	319.408	2	159.704	29.53	.0005
Time of test	256.266	1	256.266	47.39	.0005
Color x Violence	3.750	1	3.750	*	
Color x Grade	15.208	2	7.604	1.41	
Color x Time	1.350	1	1.350	*	
Violence x Grade	39.475	2	19.737	3.65	.028
Violence x Time	2.816	1	2.816	*	
Grade x Time	.908	2	.454	*	
Color x Violence x Grade	1.675	2	.837	*	
Color x Violence x Time	9.600	1	9.600	1.78	
Color x Grade x Time	8.575	2	4.287	*	
Violence x Grade x Time	3.108	2	1.554	*	
Color x Violence x Grade x Time	13.575	2	6.787	1.25	
Within groups	1167.999	216	5.407		
Total	1847.983	239			

*Value of F less than 1.00

Table C-8: Central auditory items

Source of Variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	1.837	1	1.837	*	
Violence level	.337	1	.337	*	
Grade	164.558	2	82.279	33.97	.0005
Time of test	78.204	1	78.204	32.29	.0005
Color x Violence	1.504	1	1.504	*	
Color x Grade	4.525	2	2.262	*	
Color x Time	.037	1	.037	*	
Violence x Grade	15.075	2	7.537	3.11	.046
Violence x Time	1.204	1	1.204	*	
Grade x Time	3.658	2	1.829	*	
Color x Violence x Grade	3.408	2	1.704	*	
Color x Violence x Time	2.204	1	2.204	*	
Color x Grade x Time	4.225	2	2.112	*	
Violence x Grade x Time	8.508	2	4.254	1.76	
Color x Violence x Grade x Time	3.508	2	1.754	*	
Within groups	528.099	216	2.421		
Total	815.895	239			

*Value of F less than 1.00

Table C-9: Central visual items

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	11.704	1	11.704	7.83	.006
Violence level	.337	1	.337	*	
Grade	27.075	2	13.537	9.08	.0005
Time of test	51.337	1	51.337	34.34	.0005
Color x Violence	.504	1	.504	*	
Color x Grade	3.358	2	1.679	1.12	
Color x Time	1.837	1	1.837	1.22	
Violence x Grade	6.025	2	3.012	2.02	
Violence x Time	.337	1	.337	*	
Grade x Time	3.475	2	1.737	1.16	
Color x Violence x Grade	1.308	2	.654	*	
Color x Violence x Time	2.604	1	2.604	1.74	
Color x Grade x Time	2.325	2	1.162	*	
Violence x Grade x Time	2.925	2	1.462	*	
Color x Violence x Grade x Time	3.908	2	1.954	1.30	
Within groups	322.900	216	1.495		
Total	441.962	239			

*Value of F less than 1.00

Table C-10: Peripheral-central difference scores

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	2.400	1	2.400	*	
Violence level	14.016	1	14.016	2.61	
Grade	21.758	2	10.879	2.03	
Time of test	1.350	1	1.350	*	
Color x Violence	1.666	1	1.666	*	
Color x Grade	1.575	2	.787	*	
Color x Time	.066	1	.066	*	
Violence x Grade	8.808	2	4.404	*	
Violence x Time	3.750	1	3.750	*	
Grade x Time	3.325	2	1.662	*	
Color x Violence x Grade	24.858	2	12.429	2.32	
Color x Violence x Time	11.266	1	11.266	2.09	
Color x Grade x Time	8.308	2	4.154	*	
Violence x Grade x Time	16.975	2	8.487	1.58	
Color x Violence x Grade x Time	22.258	2	11.129	2.07	
Within groups	1159.599	216	5.368		
Total	1301.983	239			

*Value of F less than 1.00

Table C-11: PA-CA difference scores

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	5.400	1	5.400	1.78	
Violence level	1.350	1	1.350	*	
Grade	31.008	2	15.504	5.11	.007
Time of test	8.816	1	8.816	2.91	
Color x Violence	4.266	1	4.266	1.41	
Color x Grade	.325	2	.162	*	
Color x Time	1.666	1	1.666	*	
Violence x Grade	12.775	2	6.387	2.11	
Violence x Time	.150	1	.150	*	
Grade x Time	3.558	2	1.779	*	
Color x Violence x Grade	5.158	2	2.579	*	
Color x Violence x Time	.066	1	.066	*	
Color x Grade x Time	8.608	2	4.304	1.42	
Violence x Grade x Time	1.225	2	.612	*	
Color x Violence x Grade x Time	20.008	2	10.004	3.30	.04
Within groups	654.800	216	3.031		
Total	759.183	239			

*Value of F less than 1.00

Table C-12: PV-CV difference scores

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	15.000	1	15.000	8.278	.004
Violence level	6.666	1	6.666	3.68	
Grade	2.858	2	1.429	*	
Time of test	3.266	1	3.266	1.80	
Color x Violence	11.266	1	11.266	6.21	.013
Color x Grade	3.175	2	1.587	*	
Color x Time	1.066	1	1.066	*	
Violence x Grade	.408	2	.204	*	
Violence x Time	5.400	1	5.400	2.98	
Grade x Time	1.108	2	.554	*	
Color x Violence x Grade	13.358	2	6.679	3.69	.027
Color x Violence x Time	13.066	1	13.066	7.21	.008
Color x Grade x Time	.058	2	.029	*	
Violence x Grade x Time	11.425	2	5.712	3.15	.045
Color x Violence x Grade x Time	.208	2	.104	*	
Within groups	391.400	216	1.812		
Total	479.733	239			

*Value of F less than 1.00

Table C-13: Picture ranking task

Source of variance	Sum of squares	d.f.	Mean square	F	P
Color-B/W	4.816	1	4.816	*	
Violence level	109.350	1	109.350	*	
Grade	8111.633	2	4055.816	20.86	.0005
Time of test	14.016	1	14.016	*	
Color x Violence	.016	1	.016	*	
Color x Grade	361.233	2	180.616	*	
Color x Time	70.416	1	70.416	*	
Violence x Grade	240.100	2	120.050	*	
Violence x Time	36.816	1	36.816	*	
Grade x Time	.633	2	.316	*	
Color x Violence x Grade	318.633	2	159.316	*	
Color x Violence x Time	16.016	1	16.016	*	
Color x Grade x Time	64.012	2	32.006	*	
Violence x Grade x Time	42.433	2	21.216	*	
Color x Violence x Grade x Time	164.033	2	82.016	*	
Within groups	42023.021	216	194.551		
Total		239			

*Value of F less than 1.00

Studies in Film- and Television-Mediated Arousal and Aggression: A Progress Report

Percy H. Tannenbaum

Graduate School of Public Policy and Institute of Human Learning
University of California, Berkeley

This paper is a progress report on a continuing program of research that was under way prior to the formation of the Surgeon General's Scientific Advisory Committee on Television and Social Behavior and will continue after that committee makes its report. It is being prepared at the midpoint of a two-year grant period, when some of the methodological problems are just being ironed out and others are still to be accommodated. Few definitive statements can be made at this stage—we have simply not progressed far enough in understanding the general phenomenon under investigation, and the work on this grant in particular—but it is possible that our results to date can contribute to the Committee's deliberations and ultimate recommendations.

The focal point of the research program is a theoretical model which posits that many communication messages—especially those of dramatic impact featuring stimuli to which we have been emotionally conditioned, such as erotic or violent materials—can evoke varying degrees of generalized emotional arousal. The model contends, moreover, that such arousal has drive potential and is thereby capable of increasing the degree of subsequent behavior an individual may be called upon to perform while a significant residual of the arousal still persists within him. As such, the model accepts the demonstration of an instigational effect of television and film messages featuring relatively high levels of aggressive content, such as that repeatedly found in experiments by Berkowitz and his associates (cf. Berkowitz, Corwin, and Heironimus, 1963; Berkowitz and Rawlings, 1963; Berkowitz, 1965; Berkowitz and Geen, 1966, 1967; Geen and Berkowitz, 1966, 1967; Berkowitz, 1969, 1970). But it tends to attribute such consequences more to the level of arousal elicited by the stimulus materials than to the aggressive content cues, as such, in the message. However, the two postulated mechanisms—the aggressive cue model Berkowitz (1969, 1970) tends to emphasize and the emotional arousal model proposed here—are not logically incompatible nor mutually exclusive, and the higher levels of aggressive behavior observed in many studies may be a result of both mechanisms operating independently or in interaction with one another.

The main purpose of the research has been to examine the feasibility of the basic theoretical model, particularly in terms of its implications for understanding the effects of violent materials in contemporary television and film fare. I make no great claims to such a theoretical stance, nor do I feel particularly committed to it. But I do believe it is a plausible model for the investigation of certain effects of communications—one that merits serious consideration and hence systematic testing and exploration. If at all valid, it may be especially appropriate to the violence issue, where arousal has been thought to play a significant, if rather ill-defined, role (Bandura and Walters, 1963; Berkowitz, 1969; Feshbach and Singer, 1971). By offering an alternative explanation for the findings of much of the earlier experimental research in this area, it also suggests an agenda for much of the forthcoming research—and possibly, if we dare look that far ahead, for public policy decisions relating to the presence of such content in our mass media of communication.

Lest we make a premature leap, the model must first be subjected to appropriate testing for feasibility and competence, before total validity can be assessed. Such testing has occupied a major portion of our attention to date, and we are still not fully certain of the model's adequacy. As with other such formulations, it does not lend itself readily to direct and definitive testing. The model does, however, lead to a number of related implications that are amenable to experimental investigation, at least in principle, especially if problems of inadequate knowledge and unreliable technology in the assessment of the physiological concomitants of emotional arousal can be accommodated.

A second item on our research agenda deals with the role of various "presentational factors"—variables that commonly accompany the portrayal of violence in the mass media and thus provide a context for its interpretation. To the extent that they are consistently present in, say, television fare (cf. Gerbner, 1971), these factors can be said to characterize a pattern or selection or censorship—of deliberate inclusions and omissions—that helps regulate the flow of such content on the video screens. Of particular concern here are the consequences, intended or otherwise, of such contextual factors in reducing or increasing any instigational effects violent messages may produce. In keeping with the general theoretical reasoning underlying this work, concern is also directed at the effects of such treatments on the level of arousal that may be elicited.

Both sets of research undertakings—those motivated by testing various implications of the theoretical model and those stemming from the presentation factors in the media—use a similar and rather standard experimental paradigm. There are features of the methodology and procedure that require systematic examination in order to clarify such issues and accommodate the possible introductions of artifacts into the experimental findings. There turned out to be a considerable number of such procedural factors, and accordingly a good part of this report is devoted to such issues.

The basic rationale for the proposed emotional arousal model has been developed in earlier publications (Baker and Ball, 1969; Tannenbaum, 1971). The reader is invited to refer to it for a theoretical formulation of the model.

METHODOLOGICAL CONSIDERATIONS

Because a major factor motivating the present project was to test an alternative theoretical model to account for the results of earlier experimental research, the same basic methodology was employed in most studies we undertook for this research program. Before we report on studies dealing with implications of the theoretical model and related issues, it would be well first to consider the critical characteristics of this methodology.

A number of procedural and design problems are inherent in the basic methodology in this area. For one thing, it was obvious that the basic procedure used by Berkowitz and others was open to an assortment of possible artifacts and contaminations. Some of these could be met in a rather straightforward manner by introducing appropriate modifications into the experimental procedure. But it was also obvious that other methodological issues could not be disposed of so readily, since they involved more fundamental questions for which answers were not readily available. More important, some of these problems, originally arising in a purely methodological framework, raised even more basic theoretical issues that deserved study in their own right. To accommodate such theoretical concerns and to avoid the risk of making *ad hoc* procedural decisions which could vitiate the results of any subsequent research, we determined to try to accommodate these various issues through appropriate research undertakings.

Basic experimental procedure

The general procedure for most experiments followed that employed by Berkowitz and his associates (e.g., Berkowitz, 1965). It typically includes three distinct phases, with a particular effort made to maintain them as separate activities as far as the research subjects are concerned.

Phase 1 involves an encounter between the experimental subject (*S*) and another apparent subject, actually a confederate (*C*) of the experimenter (*E*). In this initial stage *S* is angered (or, when a control condition is utilized, not angered) by *C*. The most common interaction is for *S* to express opinions on a variety of current issues, with *C* indicating his apparent disagreement on most items by delivering (actually mild) electric shocks to *S*. In at least one study in this series, as an additional condition, *C* is actually helpful to *S* by delivering points convertible to cash. In almost all cases, only male *Ss* have been used, with the *C* and *E* roles also played by male students.

In *phase 2*, *S* is exposed to one of several experimental films, depending on the experimental conditions. These film or television segments are usually of relatively short duration (at times as short as two or three minutes, rarely over 15 minutes). *S* is told that he is participating in a learning experiment in which he is the teacher and *C* is the learner. He is told that *C*'s task is to respond to certain questions relating to information in the film and that *S* is being shown the film to get "some idea of what it is about." In earlier research, the typical study involved two film segments presumably varying in degree of aggressive content. More recent studies in our research program have tended to compare messages varying in degree of general arousal. The arousal is usually indexed on a set of physiological measures, either in a pretest or, in some recent cases, during the exposure phase of the main experiment.

Phase 3 is the response situation, in which S usually delivers ostensible electric shocks to *C*, as "negative reinforcement," whenever *C* makes an error on a test item. There usually are a total of 20 test items, with 12 more or less randomly assigned as incorrect and hence as trials for shock delivery. It should be noted that *S* generally has to deliver a shock on an appropriate trial (either that, or withdraw from the experiment) but that he is free to choose the intensity of the shock by pressing one of ten available buttons. It is the intensity score, computed over the 12 trials, that usually constitutes the main measure of aggressiveness, sometimes in association with other related parameters such as total duration of shocks, or number of shocks per trial.

Procedural modifications

It is obvious that such an experimental procedure is frought with opportunities for "experimental expectancy" (cf. Rosenthal, 1966, 1969) and for "confederate expectancy" contamination, in which cues are seemingly conveyed to *S* guiding him to respond in a certain way. Similarly, the procedure has been criticized on the grounds of "demand characteristic" effects (cf. Orne, 1962, 1969): *S* ostensibly infers the purpose of the experiment, or at least of the condition he is in, and performs in keeping with his hypothesis of the situation rather than as he may actually feel. If either or both of these situations obtains, one could argue that the obtained responses are largely artifacts of the experimental situation and not genuine indicators of felt aggressiveness.

While such charges have of late been directed at experimentation in social psychology, they are rarely demonstrated for a particular experimental situation—and even then with only speculative, not substantive, foundation. Furthermore, it is not always clear why the intrusion of such factors should favor the results obtained, as the criticisms generally contend. Nevertheless, the risk of contamination is there, and when they are easy enough to accommodate in the particular experimental situation, the mere possibility seems justification enough to take corrective measures. There certainly seems little reason not to do so in this case, and we have accordingly introduced appropriate modification wherever possible.

Whereas, in earlier research, all interaction between the experimenter (*E*) and the subject was on a face-to-face basis, this contact is now kept at an absolute minimum. Instructions for each phase of a given study are all delivered by tape recording, with *E* serving mainly to lead *S* from one experimental phase to the other. It is assumed that by cutting down such interaction, possible contaminations of the experimenter expectancy variety are reduced, if not totally removed.

Perhaps more important, we have eliminated the necessity of *S* and *C* encountering one another in person. This was an obvious source of possible artifact; *C* was called upon to play distinctively different roles in different experimental treatments and to maintain the precise same role within a treatment. To reduce such unwanted between- and within-condition variance, the procedure has been modified so that *C* and *S* do not actually meet (although *S* has the definite impression of *C* being in the adjacent room). All communication between the two that is called for in the experiment is carried on through a coded signal system (see Zillmann, 1969, for precise details).

This change also allowed us to use females occasionally in the *C* role, easing the task of running *Ss* through a particular experiment within a reasonable time span. To further reduce the possibility of error on *C*'s part in recording *S*'s response data, we have introduced an automatic data recording system.

In general, these and other more minor alterations have worked out quite well. If anything, they appear to make for a somewhat smoother, more uniform procedure, one that is readily adaptable to different studies and varying testing conditions (We have had, on a number of occasions, to conduct experiments outside our established laboratory quarters.)

Demand characteristic effects

These modifications may help accommodate various potential contaminations, but they probably do little to meet the objections that can be raised about demand characteristics or even "evaluation apprehension" (Rosenberg, 1965) intrusions. These are somewhat amorphous issues to begin with, and often condemnations of this type turn into cases of "throwing out the baby with the bathwater." For example, it is usually assumed that if an *S* is aware of what is "expected" of him by *E*, he will behave in accord with these expectations. Actually, of course, there is as a good a reason to believe he will behave in a manner contrary to that expectation than in accord with it (see Argyris, 1968; Stricker, 1967; Stricker, Messick, and Jackson, 1969, for the notion of the negativistic subject). Thus, while contamination may well exist, it need not be in conformity with the theoretical predictions. Its major consequence in many studies might be merely to increase the within-group variance, thus increasing the error term. Similarly, unless it can be shown that these intrusions interact with the experimental condition, it is difficult to see how they would explain the particular findings that are obtained in a given experiment.

It is nevertheless desirable to accommodate such objections if possible. In discussions among our researchers, enough suspicions were generated about such potential intrusions that we decided to conduct a study utilizing one conventional procedure and probe for such effects.

The typical means of checking on demand characteristic effects is to have *E* conduct a postexperimental interview in which *S*'s suspicions of the purpose of the study are explored. However, there is every reason to believe that if *S*'s behavior was motivated by desires to "please" *E* in some manner, he would be reluctant to reveal his hypotheses directly to *E* freely and openly. We reasoned that to the degree he harbored any doubts or suspicions he would be more prone to reveal them to a fellow subject—if not on his own then under some enticement.

After a given *S* participated in the regular experimental procedure, he was ushered to a waiting room while *E* ostensibly was obtaining the funds to reimburse him for his participation. Another subject (actually another student confederate *C*) was also in the waiting room, apparently for the same purpose. *C*'s instructions were to probe *S*'s feelings about the experiment in a gradually escalating sequence.

C first waited for *S* to volunteer some remarks on his own. If none were forthcoming, *C* inquired in a neutral manner about the experiment just completed.

If no spontaneous doubts were voiced, *C* would then ask *S* what he "thought the experiment was *really* about." *S* had, of course, several options of response: he could repeat the purpose as it was originally presented to him; he could express suspicions of a different purpose, which could be either the "real" one or another that had occurred to him; he could voice vague doubts, without being specific; or he could offer no reply at all.

If *S* still offered no hint of a true suspicion, *C* would then voice his own doubts (essentially that the study could not be about the stated purpose) and thereby invite *S* to do likewise.

In all cases, *C* would suggest the actual purpose of the experiment and see how *S* reacted. Again, *S* could agree with *C*'s apparent diagnosis, express his reservations about it, or stick to the original version.

In all, 60 *Ss* participated in this phase of the experiment. Of those, only two offered any hints of their own suspicion that came close to the actual purpose of the study. Most of the rest either repeated the purpose as it was essentially presented to them (42 of 58) or shrugged off the issue (eight). Several (three) voiced vague doubts but could not suggest a reasonable alternative explanation. Two came up with alternatives that made some sense given the conditions of the testing but were far off the mark. An additional three offered suggestions that were quite bizarre and in two cases at least laughingly gave evidence that they were "only kidding."

This could hardly be said to constitute evidence for contamination because of demand characteristics of the experimental procedure. In the fourth stage, when *C* openly voiced his own suspicion of the true purpose of the experiment, only eight *Ss* showed any sign of agreeing with him, and these included several who had voiced some suspicion earlier.

Another feature of these data deserves comment. The initial study in which these *Ss* served evolved from experimental groups of 15 *Ss* each. No matter how liberally the "suspicious" *Ss* were identified, their distribution across the four conditions was fairly equal, with no overloading in one cell or another. Under such circumstances, the contribution of their suspicions would have been to increase the error term rather than to seriously affect the differences between experimental treatments.

It was not possible to conduct such a postexperimental study for each experiment run in our series, but on the basis of these findings we feel fairly confident that the results of our studies cannot be readily accounted for by intrusions due to demand characteristic effects. While this is somewhat akin to proving the null hypothesis and hence is not fully satisfactory as a disclaimer, we at least do feel that the burden of proof should be on the accusers.

Assessment of aggressive behavior

Of the many frustrating problems that currently plague the behavioral sciences, that of measuring the dependent variable is among the foremost. Nowhere is this more apparent than in attempts to investigate aggressive behavior. Clearly, the ideal situation would be one in which *S* actually inflicts harm or pain to another individual and intends it as such (cf. Goranson, 1969). However, such a situation obviously violates ethical considerations and is generally out of the question. As a result, research in this area has resorted to a wide range of verbal assessment techniques or indirect behavioral measures. The former have run the full gamut of word association procedures, TAT and Rorschach protocols, story-completion tests, selection of preferred alternatives, attitude and value scales, etc. Investigators with a preference for more behavioral measures have used observational techniques (e.g., observing kicking, punching, and the like among nursery school children) or, as indicated above, intensity and frequency of electric shocks (which, while they are not actually delivered to another person, are presumably thought to be by the *Ss*). None of these procedures is totally satisfactory, and combinations of them are sometimes used (Feshbach and Singer, 1971), on the not unreasonable assumption that multiple measures are superior to an uncertain single index. However, to the degree that these measures intercorrelate highly, treating them in combination may merely be exaggerating the same set of insufficiencies.

The electric shock technique was employed in our program of research for two related reasons. Since part of the intention was to relate this research to a set of previous experimental studies, comparability of measures was an obvious consideration. After abandoning an earlier procedure in which *Ss* could rate a graduate student's performance in the experiment on the assumption that negative ratings could actually hurt his career, Berkowitz turned to the electric shock procedure outlined, and we followed suit. In addition, we have a marked personal preference for behavioral measures—reinforced, in this case, by results of an earlier study in which verbal and behavioral measures of aggression did not agree sufficiently, and the electric shock delivery technique, as outlined, seemed a reasonable compromise solution. It is an action which could certainly be considered aggressive by *S;* because shocks were not actually delivered to another party, it sidestepped (some might say avoided facing) the ethical question. The latter issue is also accommodated; we informed *Ss* at the very outset that the study would involve the reception and delivery of mild electric shocks, and they were invited to withdraw from the experiment at that point.

It is important to note that the administration of shocks is not spontaneous or even voluntary in this case; and hence its inherent aggressiveness can be questioned (e.g., Weiss, 1970). A critical consideration seems to be that if they are to fulfill their participation in the experiment, *Ss* must deliver shocks; one can insist that the behavior must be of a more spontaneous nature to be called aggression. Actually, this is a moot point; what is important in the experiments is not whether *S* delivers a shock but the selected intensity of the shocks, and/or their number and frequency, and these are totally voluntary on *S*'s part. Thus, it is assumed that a higher shock level reflects a more aggressive predisposition or intent. Because exactly the same procedure is used in all conditions of an experiment, any differences between these conditions are attributed to different levels of aggressiveness, especially when *Ss* serve as their own controls or are assigned at random to different experimental treatments.

Is there any basis to believe otherwise? Is it at all reasonable to think that, in the procedure employed, relatively high shock intensities do not actually represent higher levels of aggressive behavior, as such? Odd as it may appear at first glance, one could argue just the contrary—that more intense shocks represent less hostility toward *C*. After all, the apparent purpose of delivering the shocks is to give information to *C*. To help overcome inhibitions against the administration of shocks, *Ss* are told that the shocks serve as negative feedback to help *C* in a learning situation. Under such conditions, the act of administering shocks as negative reinforcement can be regarded as one of benevolence rather than aggression. Given such an interpretation, more intense shocks are to be taken as helpful, rather than hurting, responses; in a similar manner, an adult might argue that a child learns to avoid a dangerous situation more readily if he is administered more extreme punishment.

A suggested new measure. This line of reasoning may seem to be stretching the point a bit, but it does raise an issue that merits clarification and that should be addressed squarely. An opportunity to do just this occurred in a study undertaken for a somewhat different purpose (to be reported in more detail at a later point), when we came up with a measurement procedure that should provide an operational answer to this problem. In that study, *Ss* were called upon to react to *C* on each of the set of 20 test items administered—as before, to deliver a shock when the response was incorrect, but also to deliver a reward (on a similar ten-point scale, for tokens later convertible to cash) when *C*'s answer was correct. Thus, in this procedure both a mean reward and a mean punishment score are available, and these can both be used to compare between conditions of the experiment. If the reward scores do not differ between two conditions but the punishment scores do, then the condition with the greater shock score clearly leads to the more aggressive behavior. On the other hand, if both reward and punishment scores differ in the same direction, no clear imputation of aggressiveness can be made. In this sense, the reward scores serve as controls for the comparison of the punishment scores, hence providing for a more sensitive measure of aggressive behavior *per se*.

Given this rationale, a variety of ways of handling the matched reward (*R*) and punishment (*P*) data are indicated. One alternative is merely to use them as two separate variables and analyze them in the manner indicated above. Similarly, one could employ the *P* data in an analysis of covariance design, with the *R* scores serving as the covariate. Or, since both measures are taken on comparable ten-point continua, single composite scores are possible—either as R-P difference scores (in which case the higher the positive value, the less the aggressiveness), or as R/P ratio scores, or various such combinations.

It is rather surprising that nobody has bothered to develop a comparative measure earlier. It is so compatible with the existing measurement procedure that it looks like the natural thing to do. More important, of course, it allows for a measure that is more convincing as an index of aggressive behavior, and it does so in a reasonably realistic way with a minimum of risk to the subject. For these reasons, we have used this procedure in all subsequent studies and plan to continue that practice in the future.

A group testing procedure. On some occasions, circumstances have forced us to abandon the electric shock measure and to look elsewhere for a usable index of aggressiveness. These have recently occurred in cases where the experiments had to be conducted at some distance from our laboratory and where group testing, rather than the laborious and time-consuming individual testing necessitated by the shock procedure, was required.

After some pretesting, we have used a variation of Berkowitz's earlier technique (e.g., Berkowitz and Rawlings, 1963) in which *S* rated the graduate student playing the *C* role in a manner where these ratings could have a damaging effect on *C*'s career. Without getting involved in the definitional question whether this is aggression or not (since no physical harm or pain is involved, "hostility" might be a more appropriate term), it is clear that, if this is carried out in a realistic manner, *S*'s behavior can be every bit as damaging to *C* than the administration of electric shocks—probably more so. This, in turn, again raises the specter of ethical considerations (*S* can presumably experience considerable guilt and self-castigation for acting in such a hostile manner toward another person), but we reasoned that postexperimental debriefing could adequately take care of that.

In some of our recent experiments, *Ss* were tested in a classroom setting, and accordingly the above technique was adapted to that situation. *C* was introduced to the class as a practice teacher and in the angering condition acted in an insulting, gruff manner. At the end of the session, with *C* out of the room, a representative of the school administration indicated that *C* was being considered for a position in the school system and explained that while ratings of teacher superintendents and the like were available, the school system was initiating a new policy of canvassing student opinions as part of the decision process on the hiring of prospective teachers. He then distributed a rating form on which the teacher could be assessed and recommended (or not) for a position. Here again, negative ratings could definitely thwart *C*'s career chances and thus would be taken as indicators of aggressive feelings.

In another such case, this basic procedure was adapted to a situation in which students at a Swedish television training school were rated for their competence as television announcers. These students were presumably vying for several select announcing openings, and both absolute and comparative ratings were taken as indicators of degree of hostility.

Measurement of arousal

An assumption underlying this work is that the emotional arousal produced by a film or videotape segment is not specific to a particular emotional state (cf. Schachter, 1971); this assumption indicates the use of a variety of physiological measures to index changes in arousal level. We originally opted for a composite index of sympathetic activation, representing a multiplicative combination of changes in heart rate and mean blood pressure (which, in turn, is a weighted average of systolic and diastolic blood pressure, favoring the former), on the assumption that modifications of cardiovascular activity best reflected the humoral basis of the theorized arousal mechanism (cf. Zillmann, 1969). Skin temperature and respiration measures were also obtained on occasion—the former because it has a slower decay rate, the latter mainly to check for possible artifacts in rapid heart rate fluctuation, although respiration rate can serve as a physiological index in its own right. More recently, we have added that psychophysiological standby, the galvanic skin response, to our repertoire. It is probably the simplest single index to obtain and yields information not exactly redundant with cardiovascular data (e.g., Lazarus, 1966; Lacey, Kagan, Lacey, and Moss, 1963). Most of these parameters have been employed at one time or another in the assessment of psychophysiological states, and we reasoned that the use of such multiple indicators, across different studies but often within a single investigation, would yield useful additional information.

Such a procedure has been a mixed blessing. For one thing, we have had innumerable and regular equipment problems. We were fortunate to have the collaboration of most helpful and knowledgeable colleagues in this work,[1] but the technology of psychophysiological data collection is not fully developed and we have had our share of the usual problems: extensive delays in the delivery of ordered apparatus, unanticipated breakdowns and maintenance problems, difficulties in obtaining suitable personnel, etc.

Difficulties in interpretation. While the technical problems have forced some delays and rescheduling of research priorities, a more important issue is the interpretation of the data that we obtained. Based on our limited experience with such measures, we did expect substantial individual differences in arousal level, but we did not expect the marked lack of consistency among the various indices across subjects. A major factor here, as Zuckerman (1971) has recently pointed out in his review of physiological measures of sexual arousal, is that there appear to be pronounced individual differences in the specific locus of reactivity. Arousal in some *Ss* may show up only in heart rate changes, others only in GSR, some in both, and several in neither but on some additional parameters. While such specificity of response is perhaps to be expected in the case of sexual excitation, it was not anticipated with our materials to the degree obtained.

The existence of such individual differences in physiological response specificity makes it virtually impossible to meaningfully analyze such data across arrays of subjects within a group and, of course, across groups representing different conditions. This situation is further complicated by the fact that while most *Ss* are consistent within themselves across a number of related stimuli, a substantial proportion shows enough variation in such response specificity—sometimes on one parameter, sometimes on another—to raise additional doubts about the computation of even crude grouped scores. A further consideration is the known factor of differential physiological responses as a function of the activity level of the neural activity involved (Lacey et al., 1963; Obrist, 1963), coupled with the not unreasonable assumption that different *Ss* may approach the same film stimulus material in a more or less active manner.

At times—but, oddly enough, not always—the result is a mélange of data that does not lend itself to conventional analysis within and across treatments. On occasion, such crude directional change scores as upward or downward movement on *any* single measure represents the only apparent basis for analysis—a condition which makes something of a mockery of our detailed measurement procedures (not unlike the use of a cannon to swat a fly). We nevertheless remain convinced that there is method to the seeming madness, that there is a systematic nature to the pattern of changes in recorded physiological response. At this time, we lack the theory and technique for deciphering the patterns that may exist, but the search should not be abandoned at this early stage. It is clear that some sort of change occurs in most subjects during exposure to most of the stimuli we have employed most of the time. The presence of any such kind of emotional arousal effect is adequate justification for pursuing the problem further in those instances where conventional analysis seems ruled out for the present.

A possible new measure. Perhaps the major problem is that we are dealing with at least second-order variables, which are highly subject to external "noise" and inadequate assessment. The assumed mechanism for the emotional arousal syndrome is one of a generalized activation system (such as Selye [1950] suggested for physical or psychological stress) characterized by the release of steroid hormones into the blood stream; these hormones in turn serve to mobilize a variety of somatic features to accommodate the arousal in some way. Physiological responses of the type dealt with here are among this secondary set of reactions, essentially byproducts of the hormonal activity. Accordingly, assessment of emotional arousal should be more reliable if it were focused directly on the initial mechanisms. Schildkraut and Kedy (1967), among others, have advocated such a strategy, centering on the release of catecholamines, especially epinephrine. The measurement problems at this level, however, are even more formidable; not the least problems are the rapid dissipation of such steroid compounds in the blood and the fact that traces of their transformations appear in highly diluted form in the urine. Again, appropriate and reliable assessment will have to await technological innovation.

One such innovation may be in the immediate offing. We have recently been in contact with a colleague[2] who is in the process of developing an integrated system for analyzing the composition of a sample of effluents, be they emissions from industrial smokestacks, from botanical plants, or (of more interest here) in human breath. The system uses spectographic analysis to detect the components of a given sample, then uses a computer memory system storing a wide array of unique spectographic patterns to identify a given element or compound and its relative presence in the sample. The system is potentially capable of working at very high speeds and of detecting a wide range of possible elements with remarkable sensitivity (e.g., five points in a million by volume).

Its potentiality in the measurement of emotional states appears to be highly promising but is still uncertain. A major question concerns the exchangeability of molecules of the catecholamine compounds from the blood to the lungs; both optimistic and pessimistic judgments have been forthcoming from a number of sources. In either case, it is certainly worth the try, and we hope to participate in this effort in the near future when some of the appropriate hardware can be assembled. Even if the sampling of known substances is not feasible because of an impenetrable alveolar membrane, the system may be useful on a "fishing expedition" comparing the degree of presence of a wide range of possible factors in the breath (possibly numbering over 4,000) between known differences in emotional state. In one such preliminary study, for example, a heavy increase in hydrogen was detected under stress conditions. Such a hydrogen change was not particularly expected, but if it holds up with sufficient consistency, it may well lead to new indicators of emotional arousal generally—and perhaps specifically for different induced states.

There is still considerable engineering work to be done before the system is fully operational, but its promise of extreme sensitivity makes it most enticing for future activity in this area. It is a much less cumbersome system for collecting appropriate samples, and the probability of artifacts appears much reduced. It is even likely that remote data collection will be possible through the use of self-sealing tubes, thus allowing for group testing outside a fixed laboratory format. We have assumed a collaborative role in this work to date, and intend to measure it even more actively in the near future.

Effects of angering phase

Another element of the basic experimental procedure that has both methodological and theoretical significance is the role of the initial encounter, in which *S* is "angered" by *C* (usually by receiving a series of mild electric shocks). After this phase, *S* is exposed to a film stimulus, and is then given an opportunity to respond to *C*—the same person who served as his original tormentor—on the learning task.

A series of questions derives from this sequence of events, not the least of which is the necessity of the angering encounter to begin with. Most studies have employed such a procedure in order to evoke subsequent behavior in the response phase. Berkowitz (1962) suggested early in his research that the mere display of violence in the absence of such an angering condition could actually inhibit subsequent aggression. On the other hand, a number of studies, most notably those by Walters and his associates (e.g., Walters, Thomas, and Acker, 1962; Walters and Llewellyn-Thomas, 1963), found induced aggressiveness toward another person without the apparent angering, and Geen and O'Neall (1969) report somewhat similar findings. In general, however, there has been significantly higher aggression with the angering condition than without.

Apart from its influence in the final response phase, the initial encounter condition may also exert some influence on the intermediary film exposure. One possibility, as Berkowitz (1969) has more recently emphasized, is that the earlier angering may function to make the individual more sensitive (hence more reactive) to the aggressive cues in the violent film condition. Our own position has been that the procedure used effectively disengages these two phases of the experiment, and that regardless of that particular connection, the main influence of the initial angering activity is to later provide a convenient and legitimate target for aggression—but only if and when the individual is put into a situation which calls for, or allows for, such behavior. There are implications of both these (not necessarily contradictory) speculative positions that merit more direct investigation, and we have undertaken several such studies.

Effect on film exposure. Whatever influences the angering activity can have on *S* while he is watching the film segment, that influence should be lessened if the angering phase does not precede the film exposure but rather follows it. This distinction may not be a most crucial one, but it is readily amenable to testing.

The experiment included two groups who shared the three standard phases of the procedure: encounter, film exposure, and subsequent shock delivery. The only difference between them was the sequence of the first two stages. One group followed the standard sequence, with the angering preceding the exposure, while the other group followed the reverse order. Thus, if the angering condition does indeed influence *S* during this film viewing in a way that affects his subsequent aggressive behavior, the first group should show significantly higher shock intensities than the second group. However, if there is no such predisposing effect, we would expect no essential differences between the two groups, since they have everything else in common.

The results completely supported the latter prediction, indicating only a negligible difference (and this in the opposite direction) between the two groups. While this finding does not rule out any other influence the angering may have, it does suggest that its effect on the exposure phase is not so critical in determining the subsequent response.

Effect on arousal. This result was further reinforced by the findings of another study which varied the initial angering activity and examined the arousal levels produced by the same film. For many of the studies in the project, it was important to obtain a fairly "clean" assessment of the emotional arousal produced by a film itself. A significant question then arose as to whether the antecedent angering treatment would indeed make *S* more sensitive to the film contents—not specifically and necessarily in terms of the aggressive cues in the message, as Berkowitz suggests, but in influencing the general arousal level produced.

Accordingly, a study was undertaken in which *Ss* were divided into two groups varying in the initial angering treatment. In this experiment, *C* reacted to *S* during the encounter phase by delivering mild electric shocks on a prearranged schedule whenever he disagreed with *S*'s expressed opinion on a number of items relating to university activities. In the High Angering condition such shocks were awarded on 8 of 10 trials; in the Low Angering condition, on only two trials. *Ss* in both groups then were exposed to the same film stimulus while a selected set of physiological reactions were recorded.

The results indicated no differences to speak of between the groups. This was one study in which the specificity of arousal was readily apparent and hence in which data analysis is somewhat suspect. But whether the data were examined on individual arousal indices or on a set of composite measures, no substantial differences could be detected. Even when individual *Ss* were matched (for susceptibility to a specific physiological measure), higher scores were as apt to be found in the Low Angering condition as not.

A smaller third group was tested without any prior encounter at all. Although the data are less directly comparable in this case (because of slightly different subject pools, etc.), there were no significant differences from either of the other conditions.

For what they are worth, the data from both preceding studies do not suggest any major effect of the angering treatment on the film exposure. While the investigations cannot testify to no effect at all (that would be like proving the null hypothesis), they do not provide any evidence for the opposite hypothesis of a specific influence.

Characteristics of the target. The nature of the possible linkage between the angering condition and the ultimate aggressive behavior is of somewhat more intimate interest. For one thing, it has methodological significance. Accomplishing the angering encounter is a somewhat cumbersome procedure, taking considerable time and effort. If this treatment does not affect the following film viewing activity and also the subsequent response task behavior, it could readily be dispensed with, simplifying the experimental procedure considerably. As we mentioned above, some studies obtained differences in experimental treatments without including such an initial treatment at all, while other studies did not. Given such ambiguous prior findings, along with the inherent difficulty of generalizing to studies in our own research program using different experimental films and a somewhat different research procedure, we decided to try to settle this problem for ourselves in the context of our particular methodology.

In addition to this methodological consideration, a significant theoretical issue is involved in the necessity of positing an angering agent to serve as a specific target against whom aggression can later be directed. The issue is reflected in two contrasting formulations. On the one hand, there is the conception of the aroused individual motivated mainly to release his pent-up feelings (whether we wish to consider these specifically as aggressive or not) and thus achieve some measure of "catharsis." A different view tends to emphasize the properties of the aggressive target; in this view, the subject's main motivation is to see to it that his earlier tormentor gets his comeuppance and is properly punished.

One way of differentiating between these formulations is to have already angered and aroused *Ss* to aggress against either the original angering agent (*C*) or another party (*O*) with whom *S* has had no interaction. Under the drive-reduction theory, there should not be much difference in the intensity of shocks delivered, since the impetus here is to release aroused feelings, and the specific target is not of particular significance.

The opposite prediction is suggested by the target-characteristic approach: shocks directed toward *C* should be stronger than those toward *O*, since *S* wants *C* to be punished but harbors no particular grudge toward *O*.

A similar contrast can be obtained if, after engaging in a first round of aggressive behavior, *S* is then given an opportunity to shock *C*. In general, keeping with the catharsis type of thinking, the prior activity should serve to diminish the subsequent aggression (*S* having presumably already relieved part of his drive). A special case is the situation in which *C* has already received punishment but by another party, and in which now a *S* whom he had angered earlier has the opportunity to deliver his own punishment. Again, the catharsis model would predict a lowering of arousal (e.g., Hokanson and Shetler, 1961) and hence less harm. The legitimate target model is not so clear on this point: it can be argued that the fact that *C* has already received punishment should diminish *S*'s desire to do him further harm (Bramel, Taub, and Blum, 1968; Doob, 1970), but the fact that *S* has not yet personally punished his tormentor might provide sufficient impetus for increased aggression.

These differing sets of predictions were all accommodated in a single (albeit somewhat intricate) experimental design and test procedure involving two separate shock response administrations. All *Ss* were first angered by *C* and all then viewed the same violent film. They were then divided into three groups, according to the prescribed victim for the first set of shocking responses. One group was given the impression they were shocking *C*, the actual initial tormentor (Tormentor group). A second group was under the impression that the target was a different individual who had delivered very few mild shocks to a different subject in the initial encounter (Nice Other group). The third group was informed their target was also another person, but one who had initially behaved like *C* by delivering particularly frequent shocks to a different subject (Nasty Other group). The two "other target" groups were included to determine if the apparent nature of the other person made a difference.

The results on this first testing session are clear and unequivocal in their support of the target characteristic interpretation. The mean shock intensity for the Tormentor group (6.29) is significantly ($p < .05$) higher than that in the conditions when another person serves as the target (combined Other Target mean = 5.22). The fact that when the other target is a "bad guy" (5.48) he gets somewhat higher though not quite significant ($.10 > p > .05$) shocks than when he is a "good guy" (4.95) further buttresses the case for the target characteristic approach; an apparently more "deserving" target gets punished somewhat more.

The second response situation represented another dimension of target characteristics. All *Ss* were given the opportunity to shock their common original nemesis, *C*, but under three different conditions of information about *C*'s experience during the first response trials.

Subjects originally in the Tormentor target group were told that the electric shock apparatus had been faulty and that *C* had not been properly shocked, so the procedure had to be repeated "now that the apparatus was working properly." Faced with the knowledge that *C* had not received his due, he is given even stronger shocks on the second round (7.09) than originally (6.29). While the within-group difference here falls just short of significance ($p = .07$ approx.), there is certainly no abatement in the aggressive behavior due to the sheer activity of discharging arousal on the first set of responses (as a strict—perhaps overly so—interpretation of the catharsis notion would suggest).

Subjects originally in the Other Target conditions were pooled for the second set of response trials and divided into two subgroups. One was informed that *C* had not been shocked before but was now available to participate in his learning task with *S* (instead of a scheduled other subject "since the other tests are running behind schedule"). Given this opportunity to retaliate against their original nemesis, even though they had already shocked another party, *Ss* now aggressed at a level virtually the same (6.16) as the Tormentor group had originally. This represented a decidedly significant ($p < .02$) increase over the earlier intensity these same *Ss* had delivered to other individuals on the first round, and again testifies in favor of the legitimate target formulation.

In the remaining subgroup, *Ss* were told that *C* had already been shocked by someone else but this had been in error since the schedule called for *C* to undergo the learning task with that specific *S* as a teacher. Here, we again find a significant ($p < .02$) increase from the first response (5.26) to the second (6.29). The information that *C* had already been punished did not deter these *Ss* from delivering the equal of what they probably would have done in the first place (judging by the precise same level of the Tormentor target group on the initial set of response measures). Thus, these data suggest that the motivation is not only that the culprit be punished, but also that the "wronged" party have a chance to retaliate.

Taken together, these findings offer clear support for the target characteristics formulation and virtually none for the drive reduction approach. In all instances where aggression can be directed against the actually tormentor, it is on pretty much the same level and always significantly greater than when the aggression is against some other person. Similarly, there is no evidence that the first round of aggression, no matter against whom it is directed, decreases the tendency to aggress against *C* on the second opportunity.

These findings do not necessarily agree with those of previous experiments which have investigated one or another aspect of the problems subsumed in this study. But since the various studies differ in considerable detail (both among themselves and in relation to the present investigation), detailed comparisons and contrasts would not be of much assistance in clarification. For our purposes, the results of this study clearly suggest that the initial angering phase be retained within the general experimental paradigm. By supporting the target characteristics line of reasoning, these experiments underline the need for including a readily available, aggression-deserving recipient within any theoretical formulation of the instigational effects of media violence. Whatever relationships may or may not exist between the film and behavior conditions, the present results point to a significant link between the initial angering session and the consequent behavior.

INVESTIGATIONS OF THE AROUSAL MODEL

While emotional reactions and physiological responses to communication messages have occupied the attention of some investigators (e.g., Lazarus [1966] for stress reactions, and Zuckerman [1971] for review of sexual excitation), they have not generally been of paramount concern in communication research. Such responses, however, can constitute effects variables in their own right and merit investigation on that basis alone. Moreover, they may have the added (and perhaps more important) significance of influencing subsequent behavior of the individual. Such emotional arousal in the organism can serve as an activating factor, thereby facilitating and accentuating the particular behavior the individual is required to perform. On this relatively primitive level, the model does not demand any particular connection between the stimulus which elicits the arousal in the first place and the type of behavioral act that is subsequently performed. The emotional arousal is taken as being of a generalized nature, hence subject to being evoked by a variety of stimuli and being equally of service in mediating a range of behavior responses.

In such a conceptualization, there is nothing special about aggressive messages or aggressive responses. The cognitive nature of such message content is presumably relevant only insofar as it contributes to the level of arousal. Nor is the fact that the response situation calls for an aggressive act of any special significance to the model. In short, heightened arousal, no matter how it is produced, should lead to higher levels of response, no matter what kind of response behavior is called for.

Applied to the situation of violent content in films and television, this basic version of the model accepts a possible instigational effect but attributes it less to the aggressive content of such messages than to their arousal potential. It suggests that—probably through processes of early conditioning and socialization—films of bloody fighting, shooting, and other assorted mayhem tend to produce a relatively high level of generalized excitation, at least as compared with other films used as "nonaggressive" controls in most experiments. This arousal, then, can accrue to the behavioral response the individual is induced to perform, leading to the higher levels of aggression noted in most of the experimental research to date.

One can readily incorporate cognitive factors—stemming from the characteristics of the message, the response task, and other situational components—into such an arousal model and make their role in facilitating the behavior task more explicit. Such a version of the model would be closer to Schachter's theory (1964, 1967), which postulates an interaction between cognitive and physiological factors in determining emotional behavior. According to this theory, the arousal is still generalized, but the individual uses available information to label the exitatory state and is thus more prone to behave in a manner compatible with that label. Here the nature of the message content is more critical; it can provide a ready explanation for the arousal and hence for the "apropriate" subsequent behavior. If the arousal is attributed by the individual to the aggressive content in the message, then the cognitively compatible aggressive behavior should be enhanced. If the arousal were attributed to cognitions less compatible with the aggressive behavioral task, aggression would be inhibited.

Our research in this area has attempted to examine implications of both the basic drive version and the more elaborated cognitive version of the general model, in separate studies and at times within the same experiment. We have investigated the effects of varying both the message content and the response behavior. These in turn have led to a number of issues involving related aspects of the model which also merit reporting here.

Variations in message content

An obvious implication of the basic emotional arousal formulation is that messages other than aggressive ones are capable of evoking heightened arousal and hence of strengthening subsequent aggressive behavior. In comparing two (or more) communications, regardless of the nature of the content, the one which produces more arousal should lead to more aggression, other things being the same. This suggests, then, an alternative interpretation of the earlier experiments demonstrating more aggression following the more violent film: the films used could have varied in arousal value as well as in inherent aggressiveness and hence might have confounded these two possible explanatory mechanisms.

It is important to note the extreme difficulty of obtaining communication messages which are identical in every respect except those which you deliberately vary. This is difficult enough with witten messages and becomes virtually impossible when dealing with such complex stimuli as go into the making of a film or a television message. It is an incredible goal to hope for when one has to resort to films produced for other purposes (as most of the research has), and it has also proved most difficult (if not impossible) to achieve in our own attempts to have films and videotapes tailormade for research purposes without controlling the production facilities. The result is a situation in which perfect controls are not to be found, and hence in which special care must be exercised in interpreting research findings and in attributing them to factors for which labels are easily found but which may be quite misleading. Among other things, this suggests replication-with-variation as one criterion for acceptance.

Erotic vs. aggressive messages. In one of the main studies in this project (and the one that really launched the entire effort), an attempt was made to deconfound arousal level and aggressive content. Since a completely orthogonal design proved infeasible—largely because a highly aggressive film is, virtually by definition, also highly arousing—the study involved a minimal comparison between two films differing in one direction on the arousal factor but in the opposite direction on aggressiveness. On the basis of pretesting (using the physiological sympathetic activity measure for arousal and ratings of each film for aggressive content), two such films were selected: a specially produced erotic film, and a prizefight scene from *Body and Soul*—the former being significantly more arousing but judged significantly less aggressive than the latter. To allow for baseline comparisons and to accommodate a possible interpretation of the findings based on a symbolic catharsis model (Feshbach, 1961), a third neutral film, lower in both arousal and aggression, was included.

The experiment has been described elsewhere (Zillmann, 1969; Tannenbaum, 1971), and these sources may be consulted for the details about selection of films, method and procedure, etc. Suffice it for our purposes here to report that all *Ss* were first angered by *C* in the manner already described, were then assigned at random to the three film conditions, and then responded by delivering shocks to *C* in the learning task situation. Mean shock intensities, computed over 12 trials, constituted the main dependent variable.

The results were clear in their support for the arousal model. *Ss* exposed to the less aggressive but more arousing erotic film treatment (5.07) were significantly ($p < .05$) more aggressive than with the more aggressive but less arousing prizefight film (3.95). That this finding was not due to a lowering of intensity in the aggressive film condition because of catharsis is attested to by the fact that the neutral film group was significantly lower in intensity than both experimental conditions.

Aggressive plus erotic content. Since eroticism and aggression can often be presented as going hand in hand, it is appropriate to inquire about the differential effects of such a condition with one in which only one of the components (in this case, the erotic) is present. It was expected that the combination, probably higher in arousal value (although this was not ascertained at the time), would lead to more aggressive behavior.

The two elements were treated separately by having a single film with a common visual component but varying in audio input. The film itself featured a mildly sensuous scene of a young lady presumably awaiting her lover. One group saw only this version. For a second group, an audio portion was added, but one that was essentially redundant to the erotic nature of the visual portion. In a third version a more aggressive sound track accompanied the video portion, as the girl contemplated various ways of murdering her would-be lover for past transgressions. (See Tannenbaum, 1971, for more complete details).

The results indicated, as expected, no real difference in subsequent aggressive behavior between the essentially redudant film only and the erotic audio treatments, but more pronounced aggression after exposure to the aggressive version. Whether this is due to its increased level of arousal stemming from combinational nature or only to its mere introduction of aggressiveness absent in the other condition is difficult to determine, but we do have some *post hoc* evidence that the erotic-plus-aggressive version is generally more arousing.

Humor message. In a subsequent study, also reported in detail in Tannenbaum (1971), it was reasoned that humor could serve as an arousing stimulus. Deliberate efforts were made to locate a humorous film short that was both funny and not aggressive (criteria that were not that easy to meet). One featuring a combination of dry humor and some slapstick was selected for contrast with the aggressive boxing match selection and the neutral travel film, in the usual experimental paradigm.

While the results indicated a significantly ($p < .05$) higher aggression score for the humor film treatment as compared to the neutral condition, they also showed the humor condition to have significantly ($p < .02$) lower scores than the aggression group. Whether this latter finding was due to cognitive interaction features or to differences in arousal is, again, difficult to tell, since a *post hoc* check for arousal level of the humor film (conducted, it should be noted, at a much later date and with a different type of subject population) indicated the arousal level of the humor film also to be somewhat below that of the violent one.

Levels of aggression. The studies reported so far have involved the contrast of one type of content with another, and in some cases also of varying levels of arousal at the same time. Another means of testing the emotional arousal model is to keep content constant as much as possible but to prepare versions of the film message to represent varying degrees of arousal. Again, the prediction from the theory is that the more arousing version would lead to more extreme aggressive behavior.

In a recent study, just such a contrast was undertaken. The availability of some film footage involving both verbal and some physical aggression and shot in a variety of ways as the basis for an exercise (unfortunately never realized) of film editing effects provided the initial impetus for the study. Through repeated and judicious editing and testing (first using only sophisticated practitioners in documentary films, but later on involving less media-oriented students) three alternative versions of the basic film were arrived at. While they are not matched in content in the full sense of the term—each has some scenes not found in the others; sequence varies on many common scenes; one features decidedly more verbal aggression, while another features the physical component—the three were quite comparable when judged on a verbal rating scale of intrinsic hostility and aggressiveness.

Judgmental measures of arousal value were also obtained at different stages of the emerging versions to guide their final selection in terms of reflecting three distinctive levels of arousal. While only gross GSR and heart changes were available to check on the terminal judgment levels before *Ss* were available for testing, it was decided to proceed with the experiment and either check the arousal values more completely later or rerun the experiment completely. For what they are worth—and, again, for reasons already cited, they may not be worth that much, with only crude indices available (e.g., a difference in any one physiological parameter)—the physiological data tend to reflect the same order of arousal as did the subjective ratings.

Pending a complete analysis of these data from three test groups randomly assigned to the three versions, there is some evidence in support of the theoretical expectations. The most arousing and least arousing film treatment groups differed significantly ($p < .05$) on aggression scores. The intermediate version, however, was not significantly different from either of the extreme groups.

To what can we attribute these differences? Again, we are working with materials that are somewhat crudely differentiated (although better than in many other studies) and measures that are not much better. An attempt was made to vary arousal while keeping content more or less constant; if we accept the manipulations and their assessment as valid, the results are further evidence favoring the arousal model. However, verbal rating scores of apparent aggressiveness are not fully reliable. It is also true that the version rated as most arousing contains somewhat more physical violence. This may be offset by the vitriolic nature of the verbal aggression in the other versions as far as the ratings on the aggressiveness scale are concerned; however, it is still possible that the higher incidence of physical violence is responsible for both the higher arousal measures and, independently of the arousal *per se*, also responsible for the higher incidence of aggressive behavior.

It is worth noting in passing that this type of study clearly brings out most of the problems involved in conducting adequate investigation in this area. A proper experimental procedure and dependent variable measures are clearly not enough. We also need reliable ways of estimating the critical selection criteria. Such a study requires an artifact-free and sensitive means of assessing arousal to check on the independent variable manipulation—where there are serious doubts that such an uncontaminated assessment is possible, given the uncertain nature of the phenomenon in the first place and doubts about whether the available instrumentation can be equal to the task. We are similarly uncertain about controlling for the content factor—although since film aggressiveness is essentially a judgmental matter, what else can one do than use such (admittedly weak) rating scales? Even if these problems were adequately accommodated, there is the everpresent problem of the inherent unlikelihood of matching messages on all possibly relevant criteria. In the present case, the relative presence of physical *vs.* verbal violence was such a problem (and will be the subject of another study in our research program); there are a host of other such potential factors in almost any study of this type. A possible saving grace is that the systematic study of those factors that do crop up (like the possible physical and verbal violence difference) may illuminate just why one message is judged more aggressive than another, or is more arousing, or whatever.

Content-free messages. If the basic drive version of the arousal model is essentially correct, the purely content characteristics of films may be of relatively minor significance in the instigation of subsequent aggressive behavior, except insofar as they provide stimulation for the elicitation of arousal. If it were possible to create a communication message which would have little, if any, distinctive content (let alone aggressive content) and still be arousing, it too should lead to aggressive behavior. Thus, while the immediately preceding study attempted to hold content fairly constant and the earlier studies put messages of different content in deliberate opposition, another way of exploring the influence of arousal is to virtually eliminate the explicit content. It was reasoned that some available films which feature movement, shape, form, and color, but few actual cues of reference or measuring, could perform such a function.

To be sure, as with other forms of abstract art, such films cannot be regarded as being "content-free" in the strict sense of the term. Most people can read meanings into them; in fact, the abstract nature of such materials often stimulates a search for a referent, for something to aid in interpretation. Thus "content-ambiguous" might be a more appropriate term for such films, and it should be possible to make them appear more or less arousing or aggressive by suggesting appropriate interpretations to the viewer.

Two studies have been initiated to investigate such possible effects, both involving films produced by the National Film Board of Canada. In one, an early film by Norman McLaren, utilizing a technique of painting directly on the film itself, was tested as a possible stimulus, with and without the sound track. For its time an especially innovative film, it coordinates a series of color, shape, and movement variations with a spirited, frolicsome folk tune. When circumstances during part of the pretesting cancelled the sound track, some difference in arousal value was detected (once again, on a cruder basis than we would prefer), and two versions were thus available. Both versions showed a significant change on a before-after basis—somewhat more so for the complete original than with the music removed, although the difference in change scores between the two failed to reach statistical significance. We decided to use both versions in comparison with a third no-film condition (oddly enough, a control condition used rarely, if ever, in this area of research).

Another variation in this study involved the use of a different measure of aggressiveness. For a variety of situational and procedural reasons, it was not feasible to use the electric shock procedure with individual testing. *Ss* in each group were tested in a classroom situation where they were first "angered" by negative and disparaging treatment by an apparent practice teacher (actually an experimental accomplice). After the film exposure, *Ss* were told that the practice teacher was being considered for a position in the school system and that in addition to the usual ratings by the school authorities, recommendations were being solicited from students. Utilizing a realistically prepared questionnaire, *Ss* were thus led to believe that unfavorable ratings of the teacher could significantly affect both his probability of placement and his salary level. Thus, the more negative the ratings, the more hostile (if not actually "aggressive") the behavior.

On the key response question regarding recommendations for hiring or firing the teacher (only one of several used to measure the hostility), the results show the most negative ratings of the teacher in the group that received no film exposure. This may be explained by at least two factors. The testing for this group occurred immediately after the practice teacher's appearance (without the shift of teachers and the intervening film), and it is possible that the students' resentment was at its peak at that time. A second explanation addresses the situation in terms of the arousal model, suggesting that the angering situation may be arousing in itself. (Unfortunately, we have never tested for excitation at this stage of an experiment). Especially in this case, the arousal for this group might be higher than the two film conditions in which the material, while somewhat arousing, is also rather gay, light, and possibly distracting. This line of reasoning—strongly suggestive of a cognitive component—might also account for the finding that the film version which included the music actually yielded the least hostile responses. Although the difference in response between the two film versions is not significant, it is opposite in direction to the arousal data (which were also not significantly different). Assuming this trend in the data holds up under more thorough analysis, it could be reasoned that the lighthearted nature of the music contributes even further to a perceived state of wellbeing and hence to a lessening of the hostility toward the teacher. Such an interpretation almost negates the role of arousal but is worth considering, particularly for such situations as are represented by this somewhat special testing format.

Another study utilizes a content-ambiguous film of more recent vintage by Pierre Hébert. This film has a more apparent aggressive intent, or at least one in which conflict can be attributed to the message more readily. Since pretesting of the physiological measures showed this film to be of only a moderate arousal level, attempts to make it more or less aggressive are being manipulated by having *Ss* read an ostensible film review and a statement by the filmmaker of his intention in making the film to begin with. Any differences that emerge from these conditions would presumably be due more to specific cognitive factors produced by such manipulation than to any intrinsic arousal value as such—although it may well be (and this too is being measured) that the version labeled "aggressive" may be more arousing to begin with. This study is in its early stages, and the complete results will not be available for some time to come.

Variations in response

In terms of the primitive version of the arousal model, the fact that electric shocks are administered in the response behavior merely makes aggression a special case of a more general phenomenon. At its extreme, this model holds that virtually any subsequent behavior would be affected by the existing state of heightened emotional arousal. The arousal is assumed to be a nonspecific energizing agent and, as such, ready to be of service to whatever behavior the individual is called upon to perform. Nothing in this model calls for the organism to seek release of its arousal —let alone release in a certain manner such as aggressive behavior, as a catharsis type of formulation would insist upon.

Humor response. Just as humor can be used as arousing material, it may also serve as a postcommunication response. At least, we reasoned that if a person were aroused in some way and then exposed to a brief comical presentation, his reaction to the humor would be influenced.

We originally hoped to use some appropriate behavioral measure of humor, such as intensity of laughter. However, several attempts in this direction proved inadequate, and we resorted to a rating scale technique. The same three films used in studying the comparative effects of humorous and aggressive messages were employed in this study as well (see Tannenbaum, 1971).

The findings again reflect some support for the basic drive model. Both the humorous and aggressive film stimuli tended to evoke significantly ($p < .05$) higher ratings of the subsequent comic routine than did the presumably less arousing neutral film. But the difference between the aggressive and the humorous condition was completely negligible, despite an apparent difference in arousal.

Rewarding behavior. In a study that set the stage for most of our recent work (among other things, leading to the new measure of aggression), responses which involved helping *C* rather than hurting him through electric shock were used as the behavioral measure. To the degree that film stimuli vary in their emotional arousal value, the basic theoretical model would predict more helping behavior following exposure to the more activating stimulus.

Two films were employed—the erotic film and the neutral travel films used in the original study by Zillmann (1969), on which pretesting indicated a maximal difference in arousal. The nature of the initial encounter between *S* and *C* was also varied systematically across these conditions: in one *C* behaved in a positive manner toward *S* (i.e., one that would be similar to the helping behavior called for in the subsequent response situation); in another, *C* angered *S* in the usual way (i.e., dissimilar to the response task); and in a third there was no *C-S* interaction at all. The response setup was the identical learning situation, except that this time, instead of delivering electric shocks as feedback on incorrect trials, *S* was called on to administer positive reinforcement only in the form of reward tokens on the correct trials.

The results are presented in detail in our previous report to the Commission on Obscenity and Pornography (Tannenbaum, 1971). For our purpose here, it is sufficient to note that, while there was a significant main effect between the two films favoring the basic drive version of the arousal model, there was also a significant interaction pattern which testified to the influence of cognitive similarity factors. Thus, the difference in effect between the two film treatments is largest when *C* was originally helpful to *S* (i.e., where the similarity between the initial encounter and the subsequent response are greatest), but this difference vanished almost completely in the angering encounter condition. Clearly, there is support here for both the arousal and the cognitive versions. In this case, the cognitive connection is between the encounter and response phases, without involving the film exposure in any direct way.

Rewarding and punishing responses. Having conducted some studies that used only the presumably aggressive electric shock response behavior and one which used an opposite rewarding task, a logical extension was to combine both measures in a single study. As a first step in this direction, the film exposure was kept constant for all *Ss,* with only two groups (reflecting relatively positive and negative initial *C-S* encounters) used. All *Ss,* however, responded in both rewarding and punishing trials, thus yielding scores for both behaviors for each subject and allowing for simultaneous comparisons on both types of behavior.

Subjects were randomly assigned to the two groups in terms of the nature of the original *C-S* interaction. During this phase of the experiment, *C* indicated his apparent degree of compatibility with *S*'s expressed opinion on a set of ten items reflecting various campus affairs by delivering a shock on each trial he is in disagreement. In the *Negative Encounter* condition, *S* received shocks on eight trials; in the so-called *Positive Encounter* condition, *S* still received some shock, but only on two of the ten trials. It was felt this distinction would sufficiently differentiate two levels of original interaction while keeping the nature of the task fairly constant.

All *Ss* in both groups were exposed to the same aggressive boxing film. They could then react to *C* in the usual learning context, delivering from one to ten points for each correct answer and from one to ten shock intensities for each incorrect answer, by pressing a button in one of the two rows of buttons supplied. There were a total of 20 trials, of which eight were randomly assigned as correct and 12 as incorrect; the same order of trials held for each *S*.

The results are shown in Figure 1 and indicate a fairly clearcut pattern. While, in general, rewards (*R*) are handed out at a higher level than are punishments (*P*) (overall *R* mean = 6.50, *P* mean = 4.29; $p < .02$), what is of significance here are the relative changes on both measures across the two conditions. This can be looked at in summary form in the R-P differences scores with each group. Keeping in mind that the lower the R-P value the greater the aggression, there is significantly ($p < .02$) greater aggression for the Negative Encounter condition (1.17) than for the Positive Encounter (3.25).

The locus for this difference is more clearly identified when we examine the individual means. While there is no difference to speak of in the R-scores between the two conditions, the P-score rises appreciably as we move from the Positive (3.56) to the Negative (5.02) encounter conditions, the difference being significant beyond the .05 level. This testifies to an impact of the encounter treatment exclusively on the degree of aggression—that is, the aggressive behavior and not the rewarding behavior is vulnerable to the contrast between the two encounter conditions. When *S* is not especially upset with *C* (as is presumably the case in the Positive Encounter condition) the P-score is significantly ($p < .01$) below the R-score. But as the aggression scores rise significantly and the reward scores drop somewhat (though not significantly), this difference becomes insignificant in the case where *S* was previously angered.

Such findings are not especially relevant to the issue of the emotional arousal arising from exposure to the film, since that variable was not manipulated here. The present study was actually preliminary to a more ambitious undertaking in order to first examine the sensitivity of the measuring device to the angering condition. By demonstrating a specific sensitivity on the aggressive dimension, the study sets the stage for a more sensitive measure for the assessment of aggressive behavior, as was alluded to earlier. At the same time, the results of the study provide additional testimony for the importance of the initial encounter condition as a prelude for a legitimized and available target in the response phase, and they thus support the earlier conclusions stemming from the target characteristic model. In so doing, they also serve to underline the need for including some cognitive components in the theoretical model.

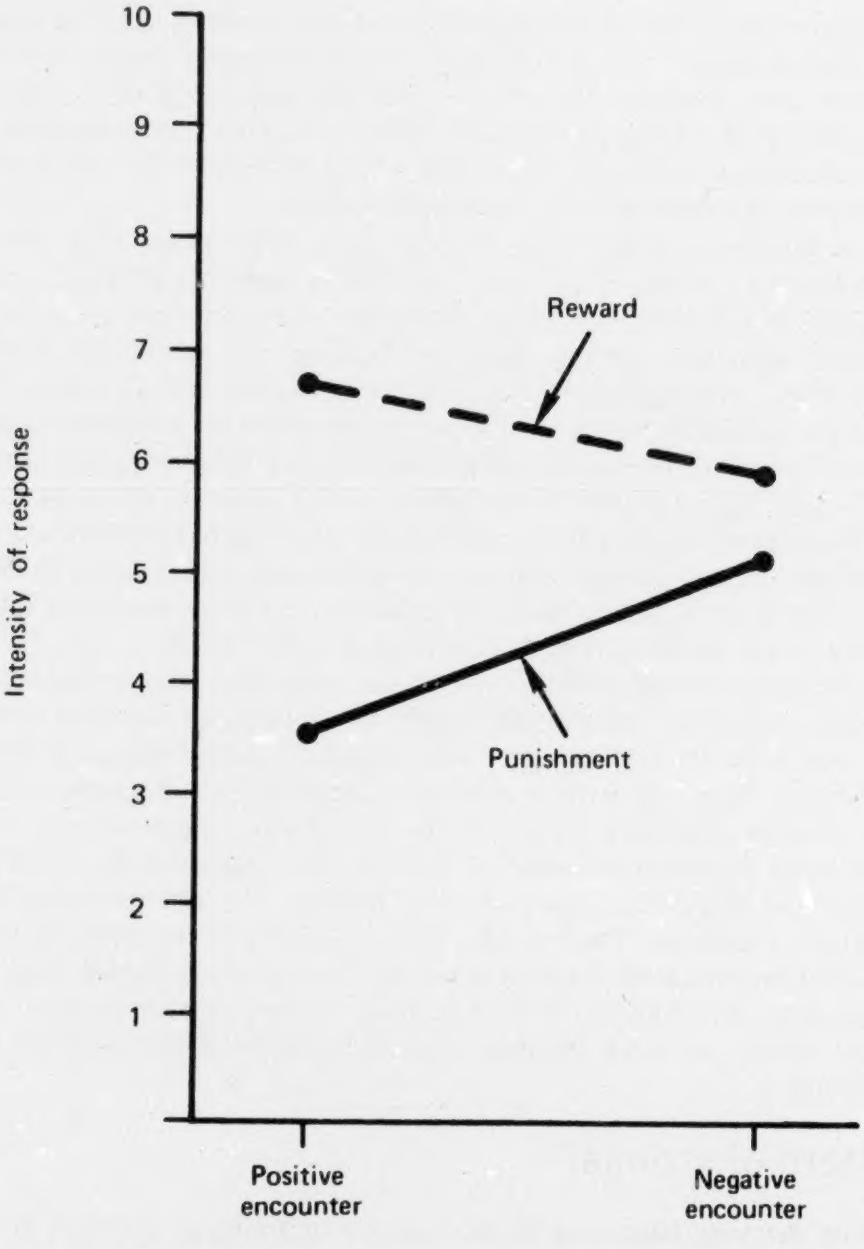


Figure 1: Reward and punishment scores as a function of a type of initial encounter

An attempt at the followup experiment referred to above has recently been conducted. In this case, variations were introduced both in the encounter condition (also as between positive and negative exchanges, but manifested in a different way) and in the arousal level of the films (aggressive *vs.* neutral), with both reward and punishment responses obtained. The results of such a study should allow for a determination of the effects of both the encounter and the more important film exposure treatments and of the interaction between them. According to the primitive arousal model, both reward and punishment scores should increase with the more arousing film and in about the same proportion. Departures from such a prediction should allow for a clearer determination of the various cognitive factors at work and possibly for the interaction between such factors and the arousal mechanism.

Ss in this investigation were students at a local junior high school which operates as one of the new "alternative curriculum" approaches along somewhat unorthodox lines. In order to accommodate the requirements of individual testing, each test session was conducted during "study hour" throughout the day. These two factors, among others, led to some adjustments in the testing procedure and were probably largely responsible for the high degree of within-group variance obtained (larger than in any other experiment conducted in this series of research studies). Because of the unorthodox curriculum and internal procedures, the school generally is thought to attract an unrepresentative—one can even call it "far out"—student body. In addition, we were forced to introduce variations in the testing routine from one *S* to another: e.g., different individuals served in the *C* role. In any case, the resulting high error variance tended to rule out any significant differences between conditions and make the findings somewhat suspect. Accordingly, we intend to replicate this study with a different sample of *Ss* in the near future. However, for what they are worth, the results of the present study indicate a more pronounced cognitive linkage effect between the initial encounter and responses situations than between the film exposure and responses situations. There is also some evidence of an interaction effect incorporating both the encounter and film variables. Again, because the various differences involved do fall short of acceptable levels of statistical results, all such findings must be considered tentative for the time being.

Pattern of arousal

From the very beginning of the active experimental research in this area, about a decade ago, a controversy has existed over the cathartic *vs.* the instigational effects of observed aggression. In accord with his theory that fantasy materials (such as the violence to be found in the mass media) provide a symbolic outlet, Feshbach (1961) reported a cathartic effect and continues to do so to this day (e.g., Feshbach, 1969; Feshbach and Singer, 1971). On the other hand, literally dozens of other studies—principally by Bandura and Walters and their colleagues (e.g., Bandura, 1965; Bandura, Ross, and Ross, 1963; Bandura and Walters, 1963; Walters and Llewellyn-Thomas, 1963; Walters, Thomas, and Acker, 1962) stressing a social learning model, and by Berkowitz and his associates (e.g., Berkowitz, 1965, 1969, 1970; Berkowitz and Geen, 1966, 1967; Berkowitz and Rawlings 1963) emphasizing the role of previously learned aggressive cues—have repeatedly demonstrated in instigational effect. These different findings may be due to such nebulous factors as wish fulfillment or experimenter expectancy effects, but a good deal of the variance can no doubt be attributed to a host of other distinguishing factors within the experimentation situations themselves: different experimenters, in different locations and different times, with different kinds of subjects, different procedures, and different dependent variables—and, not least, different film clips to represent the aggressive stimulus.

It is this difference in the type of aggressive message that interests us here, since we felt that they might be different in terms of arousal potential as well as on content factors. Our suspicions of just such a distinction were particularly prompted by the findings of Zillmann's (1969) pretesting of a number of film segments to determine their arousal values. Included in the sample were prizefight scenes from *Body and Soul*, used by Feshbach, and from *The Champion*, used by Berkowitz. As matters turned out, the segment from *Body and Soul* was selected for Zillmann's study, but in a modified version which eliminated the film's ending. Examination of the physiological data for viewers of this film revealed an increasingly high level of arousal up to the point of what may be termed the film climax (when the protagonist apparently decides not to "throw" the fight), with a steady decrease in arousal thereafter as an apparently "happy" ending unfolds.

Thus, a possible alternative interpretation of Feshbach's (1961) findings is suggested. Instead of postulating a symbolic catharsis formulation, one could argue that a buildup of excitation followed by a marked reduction would lead to a lessening of aggressive behavior. This line of reasoning suggested at least two studies: one comparing the different versions of the *Body and Soul* film, and another attempting to manipulate more directly the developing pattern of arousal and determine its consequent effects.

Happy ending effect. The main purpose of this study was to compare the two versions of the *Body and Soul* segment—that with the happy ending originally used by Feshbach, and the one used by Zillmann which was stopped at the climactic point—in terms of their effect on subsequent aggressive behavior. We already knew that these versions differed in arousal value—a finding in accord with that of Tannenbaum and Gaer (1965) reporting a reduction of stress with satisfying endings—but did not have comparable data on the response behavior.

Since excluding the film ending also meant a time difference between the two versions, an additional copy of the film was prepared which filled in most of the temporal gap with scenes of the fight repeated from earlier segments of the film. Thus, three film conditions were available: one with a happy ending, one with no such ending but which lasted an equivalent amount of time, and another without the ending but of slightly shorter (by two or three minutes) duration.

The differences were in the predicted direction; both versions of the film without the happy ending led to higher levels of aggressive behavior than among the group with the original film. These differences, however, did not quite achieve usually acceptable levels of significance, so they must be taken as tentative.

Arousal sequence. Be that as it may, the more intriguing theoretical question is the influence of the arousal pattern created by a given message on the subsequent behavior. When a given film is terminated (with the testing following shortly thereafter), the observer is left in a particular state of arousal which has several aspects worth consideration. The one obvious aspect of the state that we have tended to feature so far is the level of arousal, which may be described in absolute terms on some set of physiological parameters or, better still, in terms of the relative change from a prefilm baseline established for each individual. (The necessity of using such controls to accommodate the vast degree of individual differences on such measures cannot be overestimated.) Another aspect is the immediate history of that level of arousal. Two individuals at the same level may have arrived there from different directions—one on the upslope, the other on the downslope. Moreover, the gradient of the particular arousal slope might also be of importance—i.e., two *Ss* may both be on the downslope, but one at a sharper gradient than the other. Behind these factors is the entire preceding pattern of arousal for the film. In some cases, it might be on a more or less constant increase, for others on a steady decrease (highly unlikely); for many film stimuli, as in the original *Body and Soul,* there is a fairly consistent increase followed by an equally consistent decrease; for still others, the typical pattern might be one of oscillation, and so on.

As we have already intimated, it can be expected that each of these different patterns, slopes, gradients, and the like can exert an influence on the arousal experience quite apart from the effects of the absolute arousal level as such. In keeping with the general theoretical model, it would be expected that such differences that may be introduced would be reflected in differences in the subsequent behavior. Accordingly, a study was undertaken to explore the effects of the preceding sequence of excitation on the response output.

This was easier said than done. Our original intent was to have two film segments, one showing a pattern of increasing arousal followed by a decrease, the other showing a fairly steady low level of arousal paralleling in time the increasing pattern of the first and then showing an upturn in arousal to a point where the arousal curves for the two films intersect. Since two such precise patterns were unlikely to occur, we entertained a number of alternatives that would retain the essential characteristics of two different patterns of arousal meeting at a more or less common point (the termination point for both) from different directions. After several false starts, we opted for the use of interpolated music in two substantially edited and much reduced versions of the *Body and Soul* segment. Pretesting for arousal (in the end, because of measurement problems, only GSR was used in the selection) showed that the musical interludes did not produce the desired effects (despite deliberate selection, both types of music produced relaxation) but that the two film versions without the music still retained most of the desired qualities. One version, dubbed Decreasing Arousal, exhibited a steady buildup of excitation maintained at a fairly high level, followed by the desired decline. The Rising Arousal version showed arousal building up much more gradually and still on the rise at a point in time corresponding to the first film's downgrade phase.

Three groups of subjects were employed in the experiment, which followed the established procedure; both reward and punishment scores were taken. One group saw the Decreasing Arousal version, a second saw the Rising Arousal version, and a third group saw "early" versions of the two films (one-half of the group for each film, with the data then pooled) at a point about two and one-half minutes into the film when arousal was still in the early stages of building up.

The means for the various groups are reported in Table 1 for both reward and punishment scores and for the difference between them. Again, we find the overall R-score (mean = 7.81) to exceed ($p < .001$) the P-scores (3.18), and this holds true within each film condition. Looking at the R-P difference scores, we find that the Declining Arousal film condition (6.06) stands out as being significantly ($p < .02$ in each case) less aggressive than either of the other groups. A closer examination of the individual means shows that this is not due to any difference in the reward intensities, which remain fairly constant across the film versions. Rather, the main difference is attributable to the fact that the punishments in the Declining Arousal group are significantly ($p < .05$) below that for the Rising Arousal group as well as for the Early version. The fact that no such difference is found in the reward measure is probably best explained by a selectivity of effect deriving from the linkage between the initial encounter phase and the response activity. Since the original encounter in this study was only of the angering variety, the effect was localized to the aggression behavior.

It was impossible to fully control for absolute level of arousal of the films in this experiment. While the differences in level are not significant statistically, they leave enough doubt that we must take the evidence that a downsloping pattern evokes less aggression than an upsloping one as tentative until further data are available. What is impressive in this regard is that the condition with the highest absolute value at the point of termination is the one we labeled Decreasing Arousal, and it is on this condition that we get the least punishment—the opposite of what we would expect on the basis of arousal level alone.

Clearly, because of a number of impurities—incomplete assessment of arousal value, absence of absolute level equality, etc.—these findings

Table 1: Mean reward and punishment intensities

	By pattern of arousal		
Type of response	Early arousal	Rising arousal	Declining arousal
Reward	7.75_a	7.42_a	8.27_a
Punishment	3.51_b	3.82_b	2.21_c

Note: Means with the same alphabetical subscript are not significantly different from one another at the .05 level by Newman-Keuls test.

should be taken with more than the proverbial grain of salt. But they are supportive enough of the notion that arousal pattern influences the response to encourage further detailed activity on this issue. Whether it will be possible to vary the individual parameters of the arousal curve while controlling for all others is problematical, but some such manipulations should be attempted. The concept of emotional arousal is much too amorphous at this stage, and it is only by more systematic study of the various parameters that we can reduce the current degree of uncertainty.

Repeated exposures

The conditions under which emotional arousal is elicited as part of the set of reactions to a communication message are largely unknown. One can speak of certain types of stimuli having intrinsic, possibly universal, arousal value, as if they were innate features of the human organism, but this is obviously a debatable point. Certainly, one can point to patterns of socialization and acculturation which would also account for consistencies in such responses over wide ranges of population. However, there are individual differences, both in kind and in degree of excitation. In either case, a message may contain a set of cues that can trigger off a pattern of arousal within the individual, and these cues can interact with one another to help shape that pattern further.

This would seem to imply that, however they are acquired, such cues should have the same persistent arousal elicitation properties whenever they are presented. What then would happen if the same set of cues were repeatedly presented in sequence to the individual? If the above statement is accepted at face value we might expect essentially the same level of reaction to each of the replicated presentations. One can also conceive of the possibility of increments of excitation accruing across exposures, thus resulting in a steady increase in arousal. On the other hand, a common expectation—bolstered by some physiological response data (cf. Lazarus, 1966)—is that habituation will cause the reactions to diminish in intensity across replications. The mechanism for such an effect is unclear—it may be due to built-in limitations of the physiological reaction system or to some cognitive "tuning out" such as boredom, but that people become inured to the same stimulus is widely accepted.

All this becomes more than an academic question when applied to the repeated presentation of instances of violence on television, where repetition of aggressive acts, themes, and situations is quite common, at least within blocks of time (Gerbner, 1971). What is the cumulative effect of such a steady diet—often of the same setting, format, and specific cues—with relatively little variation? A number of related factors, of course, must be considered in attempting to formulate research to answer such a question: the specific degree of variation, time interval between exposures, contextual factors, and so on. A reasonable procedure was to start with the most redundant situation of replications of the exact same message and proceed from there. This choice also had some methodological implications, since subject availability is a continual problem and the use of the same *Ss* with the same material could be of considerable advantage in our research program.

Most of our *Ss* could be tested on three successive days, with some *Ss* tested a fourth time four days after the third trial. All *Ss* were shown two brief (approximately one and one-half minutes each) film segments of two violent situations (a prize fight in a boxing ring and a police-robbery gang shootout) on the first trial. On the subsequent trials, a given *S* was exposed to a full repetition of one of the film clips and a series of five still photographs taken from the other stimulus. Assignment of *Ss* to the two film conditions on the second or third trial was on a randomized basis, as was the order of presentation within a trial. A total of 18 *Ss* were available for the first three trials, and seven were also tested in a fourth replication (where only one film was presented). Several physiological measures were taken, the main one for our analysis being the sympathetic activation index combining pre-post exposure differences in heart rate and systolic and diastolic blood pressure.

The initial data analysis, across all *Ss* at one time, indicated that the major trend was one of habituation with a decrease in intensity from one trial to the next. Between-trial differences were definitely not significant, however. When these same data were examined separately for each *S*, this overall finding proved to be a function of a combination of three quite different groupings of subjects according to their change pattern. Nine of the 18 *Ss* exhibited the above patterns of a steady decline in intensity which held over successive trials, across both films, and, to a slightly lesser extent, whether the film clip or the photographic cue were used. An additional five *Ss* also showed such a consistent pattern, but in precisely the reverse direction, tending to increase with successive trials. The remaining four *Ss* showed a flatter change rate between trials on a more erratic pattern, tending to increase or decrease slightly from one trial to the next.

Thus, all three possible outcomes discussed at the outset of this section were uncovered in this one study. While habituation pattern dominated somewhat and hence influenced the overall effect (which was not statistically significant) more than the others, the important finding was that the *Ss* allocated themselves into three distinctive patterns, each with its own internal consistency. Under such circumstances, it does not appear very useful to refer to a general pattern of arousal over such repetitions; it seems important to recognize that individuals develop different reaction styles and means of coping with such repeated exposures.

While in general there is more emotional arousal in response to the short film segment than in response to the selected photographs, the difference is not at all significant. By themselves, the latter are almost as potent in reinstating the initial arousal as is the full message. It should be remembered, however, that the "full message" in this case was only a short segment without an explicit context, and that using such selected cues with a larger message featuring a full story line may be quite a different matter.

The one difference that reached near-significance ($p = .06$) within each of the three subgroups was the comparison between the two films. The gunfight scenes tended to evoke more arousal than did the prizefight. How much of this difference is due to these particular selections and how much can be generalized is open to question, of course, but the finding does invite speculations (which we hope eventually to translate into a research project) about the possible differential effects of (a) the use of weapons and, of more intriguing interest, (b) the notion that structured, mutually-agreed-upon conflict, such as a prizefight fought according to prescribed rules, is less arousing than a virtually rule-less, desperate (and fatal, as it turned out) gun battle. The "legalized violence" in many sports is generally regarded as less aggressive than provoked and instrumental violence, but that it may be any less arousing is obviously of interest as well.

DISCUSSION

As we stated at the very outset, this is a progress report in the middle of a continuing program of research, and it should be taken as such. We have presented the findings of a dozen or so assorted studies relating to the general issue of emotional arousal resulting from certain messages and its possible role in mediating subsequent behavior. There is a temptation—perhaps an expectation—to attempt an overall verdict of the theoretical model on the basis of such evidence, but it is one that should be resisted at this stage of the work. No single study in this series is crucial enough to be an adequate test of the theory; indeed, it is far from certain that such a test can be designed, given the somewhat amorphous nature of the issue of arousal to begin with and the obvious methodological problems. This is a situation far from unique in the behavioral sciences; the best one can do under such circumstances is to attempt to assay certain implications of the model and see how the cumulative evidence stacks up. While such undertakings cannot "prove" a theory, they can certainly constitute guides for rejecting it totally and for modifying it.

Our evaluation of the model one year ago (Tannenbaum, 1971) in light of the findings to that point was that it "fared fairly well—at the very least, it remains a plausible formulation. . . ." On the basis of the additional investigations since that time, the same general statement would appear to hold. It applies less well to the primitive drive version of the model, which focuses only on the arousal factor to the disregard of any cognitive components. There is clearly evidence in our research, as well as in earlier studies, of the role of a variety of cognitive influences. Taken together, the results strongly suggest a model incorporating both emotional arousal and cognitive components. To reject one component or the other would appear foolhardy at this point. What is obviously of significance is the interaction between arousing and cognitive factors, although it is premature to specify the particular nature of that arousal.

A pure arousal drive formulation would be hard-pressed to account for the findings of a considerable number of earlier studies—e.g., the various suggestions of cue similarity between the film and response situations (Berkowitz, 1965; Berkowitz and Geen, 1966, 1967; Geen and Berkowitz, 1966) or the results when the film violence is presented in a more justified manner (e.g., Berkowitz and Rawlings, 1963; Hoyt, 1967). A particularly puzzling case is presented by the many studies Berkowitz has conducted in which the control film features a competitive track race. In Zillmann's (1969) original pretesting, this message turned out to be not significantly less arousing than the experimental boxing match film, leading Berkowitz (1970) to argue that it thus serves as an appropriate control for the arousal component. The difficulty is that the track film is also not judged to be significantly less aggressive than the boxing film, which would imply that the eliciting cue explanation is also insufficient. Yet *Ss* exposed to the track film have exhibited less aggressive behavior than those observing the prizefight, and some explanation for such data should be available.

Our own data contain ample evidence of an interaction between arousing and cognitive components. As we have commented on several occasions in the text, this shows up particularly in terms of the relationship between the initial encounter session and the reponse situation and is especially clear in those studies in which we have assessed both reward and punishment scores. There is no evidence of any effect of the various alterations of the original interaction phase on reward scores, but it is now quite clear that when this exchange is negative for the subject, he responds more aggressively later on, and selectively so.

What is needed now is additional investigation of such possible relations between features of the film exposure phase and the response situation—and, for that matter, relations with the first phase as well. We have proceeded on a piecemeal basis to date but are now prepared to incorporate these various features into single experiments.

While we have accommodated some of the methodological difficulties in this area of research, we continue to be faced with others which, unless somewhat accommodated, will leave any findings that emerge from such experiments with a continual element of doubt. A particular case in point is the issue of the nature of physiological arousal and its assessment across a sample of individuals. The sophistication of our equipment is still inadequate to meet some of these problems—one reason why we are so fascinated with the possibilities represented by the breath analysis technique.

The other major source of possible artifact stems from inadequate means of matching sets of film and videotape messages for factors other than those we wish to deliberately vary. We have already referred to the inherent difficulty (if not impossibility) of having fully matched messages of this kind, which further emphasizes the need to have materials produced specifically for our research purposes. Some such materials have been produced, and arrangements have been made for others of a more ambitious nature, but there is still a gap between the research specifications, on the one hand, and the ability of the production unit to meet those specifications, on the other. We had made arrangements with several production agencies abroad to prepare such materials, but personnel changes and other administrative problems within the agencies caused a prolonged delay in the production of certain key materials. In one case, these internal matters seem to have been straightened out. While we are still somewhat guarded against such promises being fully realized, the need appears acute enough to warrant further efforts in this direction.

A second emphasis of our overall program of research has been the possible influence of certain variables that usually accompany the presentation of violence in standard motion picture and television fare. A number of studies have been conducted dealing with such features as the justification of violence, the effect of showing the negative consequences of depicted aggression, and the consequences of selected deletions of explicit materials from the message. Most of our activity in this area has already been reported (Tannenbaum, 1971), and we will refrain from repeating that information here. However, this general area of activity will become the principle focus of the research project during the coming year. Two additional censorship studies have already been initiated, and two others have been designed for the near future. In addition, studies of such issues as the relative effects of news *vs.* fictional drama presentation, physical *vs.* verbal aggression, and integrated *vs.* more irrelevant aggression in dramatic presentations are on our research agenda. We reasoned that it was wiser to try to solve as best we could the various methodological problems and to test the implications of the arousal model before we embarked on a large number of studies dealing with such presentational factors. While we have not resolved either of these preliminary issues fully, our time schedule is such that investigations in this area will have to proceed in the relatively near future.

FOOTNOTES

1. First, Dr. Aaron Katcher at the University of Pennsylvania, where much of our research originated, and then Dr. Curtis Hardyck at the University of California, Berkeley, where most of the subsequent research has been conducted. Our collaboration with Dr. Hardyck has allowed for the setting up of a rather well-equipped psychophysiological laboratory, complete with eight-channel tape and graphic data recording systems, an analogue-to-digital conversion unit, and a data reduction and analysis computing system capable of on-line data processing.
2. Dr. Norman Milleron, at the Lawrence Radiation Laboratory, Berkeley.

REFERENCES

Argyris, C. Some intended consequences of rigorousness in psychological research. *Psychological Bulletin,* 1968, **70**, 185-97.

Baker, R. K., and Ball, S. J. *Mass media and violence.* Washington, D.C.: U.S. Government Printing Office, 1969.

Bandura, A. Influence of models' reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology,* 1965, **1**, 589-95. (a)

Bandura, A., Ross, D., and Ross, S. A. Vicarious reinforcement and imitative learning. *Journal of Abnormal and Social Psychology,* 1963, **67**, 601-07.

Bandura, A., and Walters, R. H. *Social learning and personality development.* New York: Holt, 1963.

Berkowitz, L. *Aggression: A social psychological analysis.* New York: McGraw-Hill, 1962.

Berkowitz, L. The concept of aggressive drive: some additional considerations. In Berkowitz, L. (Ed.) *Advances in experimental social psychology,* Vol. *2.* New York: Academic Press, 1965.

Berkowitz, L. The frustration-aggression hypothesis revisited. In Berkowitz, L. (Ed.) *Roots of aggression: a re-examination of the frustration-aggression hypothesis.* New York: Atherton Press, 1969.

Berkowitz, L. The contagion of violence: an S-R mediational analysis of some effects of observed aggression. Paper presented at Nebraska Symposium on Motivation, March, 1970.

Berkowitz, L., Corwin, R., and Heironimus, M. Film violence and subsequent aggressive tendencies. *Public Opinion Quarterly*, 1963, **27**, 217-29.

Berkowitz, L., and Geen, R.G. Film violence and the cue properties of available targets. *Journal of Personality and Social Psychology*, 1966, **3**, 525-30.

Berkowitz, L., and Geen, R. G. Stimulus qualities of the target of aggression: a further study. *Journal of Personality and Social Psychology*, 1967, **5**, 364-68.

Berkowitz, L., and Rawlings, E. Effects of film violence on inhibitions against subsequent aggression. *Journal of Abnormal and Social Psychology*, 1963, **66**, 405-12.

Bramel, D., Taub, B., Blum, B. An observer's reaction to the suffering of his enemy. *Journal of Personality and Social Psychology*, 1968, **8**, 384-92.

Doob, A. N. Catharsis and aggression: the effect of hurting one's enemy. *Journal of Experimental Research and Personality*, 1970, **4**, 291-96.

Feshbach, S. The stimulating versus cathartic effects of a vicarious aggressive activity. *Journal of Abnormal and Social Psychology*, 1961, **63**, 381-85.

Feshbach, S. The catharsis effect: research and another view. In Baker and Ball (Eds.) *Mass media and violence.*

Feshbach, S., and Singer, R.D. *Television and aggression*. San Francisco, California: Jossey-Bass, 1971.

Geen, R.G., and Berkowitz, L. Name-mediated aggressive cue properties. *Journal of Personality*, 1966, **34**, 456-65.

Geen, R. G. and Berkowitz, L. Some conditions facilitating the occurrences of aggression after the observation of violence. *Journal of Personality*, 1967, **35**, 666-67.

Geen, R.G., and O'Neall, E.C. Activation of cue-elicited aggression by general arousal. *Journal of Personality and Social Psychology*, 1969, **3**, 289-92.

Gerbner, G. Violence in television drama: trends and symbolic functions. *Television and Social Behavior:* a report to the Surgeon General's Scientific Advisory Committee, Vol. 1. Washington, D.C.: U.S. Government Printing Office, 1971.

Goranson, R.E. Observed violence and aggressive behavior; the effects of negative outcomes on the observed violence. Unpublished doctoral dissertation, University of Wisconsin, 1969.

Hokanson, J. E., and Shetler, S. The effect of overt aggression on physiological arousal level. *Journal of Abnormal and Social Psychology*, 1961, **63**, 446-48.

Hoyt, J. Vengeance and self-defense as justification for filmed aggression. Master's thesis, University of Wisconsin, 1967.

Koestler, A. *The short in the machine.* New York: MacMillan, 1967.

Lacey, J.I., Kagan, J., Lacey, B. and Moss, H.A. Situational determinants and behavioral correlates of automatic response patterns. In Knapp, P.J. (Ed.) *Expression of the emotions in man.* New York: International University Press, 1963.

Lazarus, R.S. *Psychological stress and the coping process.* New York: McGraw-Hill, 1966.

Obrist, P.A. Cardiovascular differentiation of sensory stimuli. *Psychosomatic Medicine,* 1963, **25** 450-59.

Orne, M. On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *American Psychologist,* 1962, 417-24.

Orne, M. Demand characteristics and the concept of quasi-controls. In Rosenthal, R. and Rosnow, R. L. (Eds.) *Aritifact in behavioral research.* New York: Academic Press, 1969.

Rosenberg, M. When dissonance fails: on eliminating evaluation apprehension from attitude measurement. *Journal of Personality and Social Psychology,* 1965, **1**, 28-42.

Rosenthal, R. *Experimental effects in behavioral research.* New York: Appleton-Century-Crofts, 1966.

Rosenthal, R. Interpersonal expectations: effects of the experimenter's hypothesis. In Rosenthal, R. and Rosnow, R. L. (Eds.) *Artifact in behavioral research.* New York: Academic Press, 1969.

Schachter, S. The interaction of cognitive and physiological determinants of emotional state. In Berkowitz, L. (Ed.) *Advances in experimental social psychology,* Vol. 1. New York: Academic Press, 1964.

Schachter, S. Cognitive effects on bodily functionary: studies of obesity and eating. In Glass, D.C. (Ed.) *Neurophysiology and emotion.* New York: Rockefeller University Press, 1967.

Schildkraut, J. J., and Kedy, S.S. Biogenic amines and emotion. *Science,* 1967, **156**, 21-30.

Selye, H. *Stress.* Montreal: Acta, 1950.

Stricker, L. The true deceiver. *Psychological Bulletin,* 1967, **68**, 13-21.

Stricker, L., Messick, S., and Jackson, D. Evaluating deception in psychological research. *Psychological Bulletin,* 1969, **71**, 343-51.

Tannenbaum, P. H. Emotional arousal as a mediator of communication effects. *Technical reports of the Commission on Obscenity and Pornography,* Vol. 8. Washington, D. C.: U.S. Government Printing Office, 1971.

Tannenbaum, P.H., and Gaer, E. Mood change as a function of stress of protagonist and degree of identification in a film-viewing situation. *Journal of Personality and Social Psychology,* 1965, **2**, 612-16.

Walters, R.H., and Llewellyn-Thomas, E. Enhancement opunitiveness by visual and audiovisual displays. *Canadian Journal of Psychology*, 1963, **17**, 244-55.

Walters, R.H., Llewellyn-Thomas, E., and Acker, C.W. Enhancement of punitive behavior by audio-visual display. *Science*, 1962, **136**, 872-73.

Weiss, W. Effects of mass media of communication. In Lindzey, G. and Aronson, E. (Eds.) *Handbook of social psychology, 5*, Boston: Addison-Westey, 1970.

Zillmann, D. Emotional arousal as a factor in communication-mediated aggressive behavior. Unpublished doctoral dissertation, University of Pennsylvania, 1969.

Zuckerman, M. Physiological measures of sexual arousal in the human. *Psychological Bulletin*, 1971, **75**, 297-329.

Television and Aggression: a Discussion

Catharsis of Aggression Among Institutionalized Boys: Fact or Artifact?

Robert M. Liebert, Michael D. Sobol, and
Emily S. Davidson

State University of New York at Stony Brook

In a recent monograph (*Television and aggression*) Feshbach and Singer (1971) have reported an attempt to examine the influence of predominantly aggressive and predominantly nonaggressive television "diets" on adolescent and preadolescent boys over a six-week period. Endeavoring to apply laboratory methods in a field setting, they drew their subjects from seven residential institutions: three private schools serving boys from upper-middle-class backgrounds and four boys' homes for mildly troubled youngsters and those whose parents are unable to care for them (p. 52). The primary data were daily behavioral ratings, provided by institutional personnel, on a behavioral check list.

In three of the boys' homes those observing the predominantly nonaggressive programs showed *more* aggression toward peers than did their aggressive diet counterparts (p. 75). In two of the homes, significant differences in the same direction were found for aggression toward authority figures (p. 76). No differences as a function of treatment were found for boys in the private schools. From these data the authors conclude that "...exposure to aggressive content in television over a six-week period does not produce an increment in aggressive behavior ... The results, in fact, indicate that witnessing aggressive TV programs reduces rather than stimulates the acting out of aggressive tendencies in certain types of boys (p. 140)."

These conclusions are at odds with the preponderance of experimental (Berkowitz, 1970; Hartmann, 1969; Walters and Thomas, 1963) and correlational (Dominick and Greenberg, 1970; 1971; Lefkowitz, Eron, and Walder, 1971; McIntyre and Teevan, 1971; McLeod, Atkin and Chaffee, 1971) research with adolescents which has been addressed to the same issue.[1] Moreover, although the investigators are duly and appropriately cautious regarding the possibility that the conclusions will form the basis for new "therapeutic" regimens (e.g., an enhanced diet of televised aggression to reduce overt acts of aggression in troubled youngsters), their remarks might well instigate some direct efforts at application.[2] It is therefore appropriate to consider the methodological flaws of the study and point out the plausibility of certain rival hypotheses that might account for their anomalous results. We have endeavored to offer such a critique in the present paper.

Virtually no investigation in the social sciences is conducted without blemish, but some problems are likely to be more important than others in legitimately challenging any particular study. We have therefore not attempted to detail all of the problems of this investigation that have come to our attention. Among those not considered further are: the investigators performed nonparametric analyses when the more potent and stable parametric forms of analysis might have been used by employing transformed scores; many trends are described as "linear" which are quite apparently curvilinear when plotted; there was a surprisingly high attrition rate for an experimental study conducted in residential institutions; and the investigators failed to assign subjects in the boys' homes to groups randomly on an individual basis although they followed this practice in the private schools. Instead, four other methodological difficulties which we judge to be particularly serious will be considered in some depth.

1. *Reliability of the primary dependent measure was not established.*

One important criterion for evaluating any investigation is that the primary dependent measure be reliable. This problem is ever-present in all types of research and with all types of dependent measures, but it is of especial concern when behavior ratings are the measures and when the raters are relatively untrained observers rather than experienced personnel.

Feshbach and Singer, who employed ratings by untrained institutional personnel (including cottage supervisors, teachers, "houseparents," and "proctors" . . . p. 56), did not adequately determine the reliability of their primary dependent measure. The effort which they do report is described below.

> The behavior rating scale was pretested on boys at the Fernald School at UCLA and, on the basis of the pretest, was revised. A reliability study conducted between two raters who, on a specific date, had differential degrees of contact with a group of fifteen boys reflected a high correlation between the raters in their rank ordering of boys although there was less agreement in the individual rating scores recorded for each child. This latter difference was to be expected inasmuch as one rater was the classroom teacher while the second was a teaching supervisor who observed the class for various periods during the day (pp. 56-60).

It is our contention that this "check" is both inadequately documented and largely irrelevant to the reliability of ratings obtained in the experiment itself. Regarding the first point, it should be noted that the term "high correlation" (among the ranks) is not supported by an actual reliability value and is thus uninterpretable; that the degree of training, if any, of the two raters is not disclosed; and that the actual correlation between the individual aggression ratings cannot be evaluated from the remark that it reflected "less agreement" than that shown by the rankings. Additionally, no evidence was collected on the possibility of *instrument decay* (i.e., whether the reliability of the ratings will hold up over time). It is clearly possible that the ratings obtained on the first day will be more reliable than those obtained after 15, 20, or 25 days (cf. Reid, 1970).

Regarding the second point, we presume that the measure was revised *after* this reliability check and thus that *the actual instrument used was never checked for reliability at all.* Likewise, it appears that the procedure described above did not involve boys from the same samples as those employed in the actual study, nor were the raters drawn from the sample of teachers, houseparents, and proctors who provided the data reported in *Television and Aggression*.

Addressing this problem indirectly, Feshbach and Singer argue that they obtained evidence which demonstrates a satisfactory degree of agreement among the raters: ". . .in ratings [of peer aggression] from sixteen of the twenty-two raters who rated both control and aggressive TV subjects, the mean for the control subjects is higher than the mean for the aggressive TV group" (p. 68). The investigators go on to point out that this split differs from chance (i.e., totally random ratings) at the .05 level. Comparisons with chance are not, however, sufficient to establish inter-rater agreement. Indeed, that 72 percent of the raters were in accord on a lenient measure in which 50 percent concordance represents random guessing is *not* impressive. Moreover, this method of analysis revealed that agreement on the ratings *did not reliably exceed chance on the measure of aggression toward authority* (p. 69).

2. *The aggressive programs may have been preferred to the nonaggressive ones, implicating an alternative explanation of the results.*

Feshbach and Singer note on several occasions that they "... were concerned about the possibility that boys might resent being assigned to the nonaggressive diet" (p. 147). The concern is certainly well-founded. If the control group showed such resentment it might very plausibly be manifested in their aggressive behavior during the course of the experiment and totally invalidate or overshadow any other influences. The problem must be looked at from several angles.

First, there is the question of whether the programs were differentially liked. Data were apparently collected prior to the beginning of the experiment which might bear on this issue. The authors note that they ". . .recognized from the very beginning of the study that boys preferred aggressive TV programs to nonaggressive programs, and we were concerned about the possibility that boys might resent being assigned to the nonaggressive diet" (p. 147).

This problem is substantiated by the fact that a sample of six aggressive programs received more favorable ratings at the outset than did six nonaggressive programs (p. 109; the difference is described as *significant* but the associated probability value is not reported). Treatment and interpretation of the problem should therefore be considered.

> On the basis of other findings and preliminary interviews before the experiment proper was undertaken, it was our impression that boys seemed to prefer TV programs with aggressive content to programs with nonaggressive content. The reports of personnel at the several institutions and our own observations during the study indicate that the difference in preference was not major and that the control subjects had the opportunity to see programs they enjoyed . . . . The subjects were asked to rate each program they watched on a six-point scale ranging from "liked it very much" to "disliked it very much." The aggressive TV group tended to rate their programs more positively than did the controls. However, the difference between the groups is small, a substantial majority of both sets of ratings falling in the positive categories. Thus 70 per cent of the control and 77 per cent of the aggressive TV programs were rated "liked it very much" or "like it fairly much" while only 15 per cent of the control and 8 per cent of the aggressive TV programs received "dislike" ratings (pp. 131-132).

We find the argument recorded in the foregoing passage to be inadequately documented and insubstantial. The statement that "the difference in preference [between aggressive and nonaggressive program ratings] was not major" is meaningless without complete quantification. The presentation of the rating data is objectionable because the data for both groups on the entire six-point scale have not been presented. Instead the reader is left to interpret an arbitrary dichotomization ("liked it very much" combined with "liked it fairly much" *vs.* "dislike") *in which 15 percent of the data is omitted entirely.* This is hardly adequate proof of the near equality of the subjects' reactions to the programs. They may (or may not) appear to diverge much more sharply if all of the data are considered and the full range of categories is employed in the analysis. Being forced to watch relatively less enjoyable programs might, in turn, have produced resentment, frustration, or otherwise instigated aggression among control subjects.

Feshbach and Singer argue that we can nonetheless discount the possibility that the change in TV diet imposed upon the control groups made them more aggressive, stating:

> . . .if the controls were frustrated because they were taken off their favorite aggressive TV programs, then the experimental effect should be maximal in, and perhaps restricted to, the high frequency group. However, boys whose prior TV viewing frequencies appear to be less were just as affected by the experimental treatment (p. 120).

This is simply not an adequate refutation of the frustration or resentment explanation of the data. First, we question the logic of asserting (implicitly) that amount of resentment or frustration increases linearly with the total amount of TV viewed prior to deprivation. It is quite possible that a boy deprived of three favorite programs feels just as resentful or frustrated as one deprived of ten favorite programs. Second, even if such reasoning were correct, the assertion itself cannot be evaluated in the absence of information on the variability of these frequencies, according to program type, prior to treatment. If the sample variability is small (i.e., as would be the case if the "low frequency" subjects watched only slightly less than the "high frequency" subjects), then the difference between these groups in resentment would also be small and not sufficient to produce differences in the amount of aggression shown within the control group.

It therefore remains plausible that resentment on the part of the control group subjects, an important rival hypothesis recognized and raised by the investigators themselves, accounts for the high levels of aggressive behavior shown by these groups at some schools.

3. *Response bias of the raters can account for the obtained results.*

One important rival hypothesis in any experiment employing judgmental ratings is the possibility that biases, beliefs, or expectations of the raters have artifactually produced the results. Certain circumstances appear particularly conducive to (but do not necessarily produce) such bias. These include the use of untrained raters, spearation in time between the observations and their actual recording by the rater, prior familiarity between the rater and the subject being rated, familiarity of the rater with the treatment which the subject is receiving (particularly when vis-a-vis the treatment received by subjects in other groups), and the availability of hypotheses or expectations which are plausible to the rater (cf. Rosenthal, 1966).

Regarding these points, it appears that the raters in the Feshbach and Singer study were untrained, that the ratings were filled out daily (p. 56) and required evaluations that were both retrospective and judgmental (cf. pp. 162-163), and that there was prior familiarity between the raters and the subjects (p. 56). Likewise, it appears that most or all of the raters must have been familiar with the treatment received by each person he was rating, as well as with the fact that some boys in the same or adjacent dwelling units were being exposed to a different "television diet."

Most important, however, is the question of expectancies and hypotheses on the part of the raters, particularly as these might bias the results in favor of the direction of the obtained outcome. At the outset, the subjects and the cottage supervisors and teachers who would later rate their behavior were told

> . . . that the study concerned the relationship between the evaluation of different types of TV programs and the personality and attitudes of the viewer. They were further told that they would be assigned to a specific set of programs and that one of the conditions for participation in the experiment was that they stick to the specified set of programs (p. 53). [As will be considered subsequently, this obviously important requirement does *not* appear to have been met.]

The foregoing cover story, together with a few elaborations, may or may not have been convincing. It is the investigators' responsibility to show that the former was the case. They state:

> Most of the boys and supervisors appeared to accept these explanations. Intensive interviewing of a sample of participants following completion of the experiment indicated that while several felt that the object of the project was to study the influence on children of exposure to aggression on television, a substantial majority accepted the explanation of the experimenters or else entertained some other hypothesis irrelevant to the main purpose of the study (p. 53).

We do not feel that there is sufficient documentation in the foregoing passage (or elsewhere in the Feshbach and Singer monograph) to eliminate the possibility of a substantial rater bias. The contents of the interview alluded to above, the number and composition of the subsample interviewed, the actual percentage of interviewees (rather than "several" and "a substantial majority") who responded in various ways, and similar details are all needed in order to evaluate the evidence. This information is not presented.

But over and above the inadequate reporting of data in the foregoing passage, the conclusion reached stretches the present writers' credulity. Specifically, the boys' behavior was rated five days a week, for a minimum of six weeks, on a 26-item rating list *of which 19 items focused on aggressive behavior.* It seems almost inconceivable to us that an individual collecting these data could fail to surmise that the study was designed to deal primarily with aggressive behavior as a function of television viewing, especially in view of the program "diets" involved.

What hypotheses might the raters then have generated about the boys' behavior? One possibility is that they generated for themselves some version of the catharsis hypothesis. According to Feshbach and Singer, this is unlikely since ". . .when, during preliminary contacts with the various institutions, opinions were volunteered concerning the influence of television on boys, the catharsis hypothesis rarely appeared" (p. 149). It is, in our judgment, not appropriate to rely on "volunteered" opinions (by an undisclosed number of persons), nor can we evaluate the description "rarely appeared" without some sort of quantification. But more important, we do not feel that the catharsis hypothesis is the only, or even the most likely, notion to be entertained by school personnel under these circumstances.

Instead, a staff member might reason as follows: "Most boys liked aggressive programs better than nonaggressive ones before the experiment began, and many watched such shows regularly. When deprived of something which they like and have previously enjoyed, they will become more irritable and thus more likely to be aggressive and rowdy."

There are two different mechanisms through which the foregoing possibility could operate. One, discussed earlier, is that the explanation is a viable account of what actually happened at some schools. But a second and equally important possibility *is that the raters simply perceived the situation this way.* If so, then, notwithstanding their personal hypotheses about the effects of television *per se,* rater bias would clearly be in the direction of perceiving the "control" group youngsters as likely to be more aggressive than those who viewed the preferred experimental programs.

4. *The experimental and control groups were treated differentially on dimensions other than the manipulation of the independent variable.*

Perhaps the most critical requirement for preserving the internal validity of an experimental study is that all groups be treated in an identical fashion except for the manipulation of the independent variable. Unfortunately, this requirement was not met in the Feshbach and Singer experiment:

> When a number of boys in the control group at three of the institutions objected very strongly because *Batman* was not on their list, they were permitted to watch *Batman.* We were very conscious of the fact that, on the whole, the aggressive diet was more attractive to the boys and we tried to minimize any frustration associated with being assigned to the control diet (pp. 55-56).

While the procedure described above may or may not have minimized frustration, it constitutes an important difference in the experience provided for the experimental and control groups at three of the institutions that confounds the treatments being studied. Moreover, this yielding by the experimenters to a "very strong objection" by the control subjects may well be construed as a reinforcement for such related actions as grumbling, complaining, breaking rules, becoming sullen, refusing tasks, acting bossy (all of which appear among the 19 aggressive behaviors which comprised the dependent measure, pp. 70-71).

Since the strong objection (and subsequent yielding by the experimenters) is mentioned for only three of the seven institutions, and since significant differences were observed between the experimental and control groups in only three of the institutions on one measure (peer aggression) and only two institutions on the other measure (aggression toward authority), we endeavored to determine whether the confounded institutions overlapped with those in which differences were obtained. If they did not, then the confound pointed out here could be discounted to some degree from a practical viewpoint. We were only able to identify two of the three institutions in which *Batman* had been demanded, from the following passage:

> In the case of one particular aggressive program, *Batman*, permission was granted to the control group boys at Institutions B and C to watch the program since many of them had requested it . . . (p. 128).

From an earlier chapter (pp. 75-76) we are able to see that these two institutions, "B" and "C," are two of the three institutions in which significant differences were obtained on the overall measure of aggression toward peers. Likewise, "B" is one of the two institutions in which significant differences were detected on the measure of aggression toward authority. We therefore submit that the so-called "experimental effect" reported in *Television and Aggression* is hopelessly confounded, since it might have resulted from the subjects having won an unreasonable demand from the experimenters in one group but not in the other, rather than from differences in the "diets" *per se.*

Conclusions

Four potentially crippling weaknesses have been identified in the Feshbach and Singer study: 1) the reliability of the behavior rating scale was not established; 2) differential liking and/or initial preferences for the two TV diets may have led to frustration or resentment and thus to greater aggression on the part of the control group; 3) rater's expectations or biases alone might have produced the results obtained; 4) differential treatment of the experimental and control groups confounded the study and the obtained effects.

We do not suggest that any one of the plausible rival hypotheses is necessarily correct. We do suggest that the study was conducted in such a manner that *no* inferences are possible and have tried to draw attention to some of the conflicting but inseparable alternatives. It appears to us that Feshbach and Singer have drawn untenable conclusions from an inadequate design. Uncritical acceptance of their findings, especially outside of the scientific community, could have results that are at best disappointing. At worst, they could be disastrous.

FOOTNOTES

1. We say at odds with the correlational literature advisedly, since *causation does imply correlation* and the evidence uniformly shows a positive association between watching aggressive television and various measures of aggressive behavior. The catharsis hypothesis implicitly "predicts" and requires a negative relationship.
2. The dust jacket of their volume states: "The findings refute several popular theories about the impact of violence on television . . . and will undoubtedly have a serious bearing on future programming."

Television and Aggression: A Reply to Liebert, Sobol and Davidson

Seymour Feshbach
University of California, Los Angeles

and

Robert D. Singer
University of California, Riverside

The book, *Television and Aggression*, reports research findings which run counter to the views of social scientists who believe that imitation and arousal are the only effects of observing aggressive sequences on television. The findings, while consistent with certain experimental data, are indeed at odds with the preponderance of laboratory studies. The diverse findings regarding the effects of aggressive television content on aggressive behavior can be reconciled by recognizing the complex, multifaceted nature of the interaction between the observer and the media content and by examining the conditions which facilitate such mediating processes as modeling and arousal, and those which facilitate catharsis

and cognitive control mechanisms. Liebert, Sobol and Davidson instead have chosen to reconcile the differences by searching for supposed "crippling weaknesses" in the Feshbach and Singer *Television and Aggression* study. In so doing, they have found weaknesses which were not there, have ignored a great deal of relevant data and have failed to cope with the theoretical challenge and opportunity posed by ostensibly conflicting data.

The authors of *Television and Aggression* recognize that there are methodological difficulties in this type of field experiment. These are discussed at length in the volume, and much of the statistical analysis is devoted to methodological checks and the investigation of alternative hypotheses. While there may have been methodological problems which have been overlooked or perhaps unduly minimized, we can show that the Liebert et al. critique is without substance.

Their introductory comments also convey theoretical and methodological orientations which we do not share. For example, their statement that "the catharsis hypothesis implicitly 'predicts' and requires a negative relationship" (between watching aggressive television and measures of aggression) is an unsophisticated interpretation whose inadequacy has been noted for some time. The relationship between observation of aggressive television and aggressive tendencies is a function of at least two factors: (1) strength of attraction to aggressive television content, (2) effects of observation of aggressive content on aggressive behavior tendencies. A very aggressive child might have a much stronger need to watch aggressive television than a nonaggressive child. At the same time, observation of aggressive television could have a cathartic effect for the very aggressive child. What the catharsis hypothesis does imply is that for individuals with strong aggressive tendencies, various types of vicarious aggressive activity, including observation of aggressive television, can serve to reduce or control aggressive impulses and acting out behavior. Since the resultant aggressive behavior of hostile and violent individuals who observe aggressive television might still be greater than that of unaggressive individuals, one could well obtain a *positive* relationship between aggressive behaviors and observation of aggressive television even though catharsis is operating. Technically, the issue has to do with the relationship between changes in drive strength or incentive and changes in habit strength or expectancy.

Another point at issue in the introduction is the use of parametric versus non-parametric analyses. Since the behavior data were markedly skewed in the direction of low aggressive scores, it would have required radical transformations, indeed, to permit parametric analyses. Nonparametric analyses of variance were clearly more appropriate and while they did not permit the isolation of curvilinear trends or a quadratic component, they had the virtue of being more conservative and more *justified*, considering the nature of the obtained data.

There are other aspects of the introduction which bear further comment (e.g., the "surprisingly" high attrition rate to which Liebert et al. refer occurred in two of the private schools where there was no experimental effect while, in the boys' homes, where the experimental effect was obtained, the attrition rate was surprisingly low), but we shall limit our subsequent comment to the four major points they judge to be particularly serious.

(1) With regard to the reliability of the behavior ratings, we agree that more evidence of its reliability could have been obtained before we went into the field with it. However, a number of empirical findings attest to its reliability. Boys initially above the median in nomination of aggression by their peers participated in significantly more aggressive incidents (as determined by the behavior ratings) than boys below the median on the peer nomination measure. This correlation provides evidence of the validity as well as reliability of the behavior ratings. Measures of aggression based on the personality inventory also were significantly related to aggression scores as determined by the behavior ratings. Finally, the behavior ratings were sufficiently reliable to yield significant experimental effects.

Liebert et al. also make several incorrect assumptions about the reliability study reported in the book. Before attempting this reliability check, several preliminary versions had been prepared. Following the reliability check, one item out of the 26 behaviors that were rated was modified to increase its clarity. The underlined statement, "the actual instrument used was never checked for reliability at all" is an unfortunate exaggeration. Moreover, there is further confusion regarding the simple descriptive observation that, while the correlation between the two raters was high, the aggression ratings of the teaching supervisor tended to be lower than those of the classroom teacher. The actual Spearman Rank Order Correlation was .88. It was not reported because we felt that it might convey a misleading impression of the reliability of the instrument as actually employed in the experimental situation. The point was that the reliability was sufficiently high to encourage its use in the experimental situation. There did not appear much risk in employing the instrument because: a) many comparable items had been used in related observation inventories of established reliability; b) the ratings of the teacher supervisor who was more removed from the children than the classroom teacher nevertheless correlated highly with the ratings of the latter; c) we had several training sessions (2 to 3) with the raters in which the instrument was discussed, examples brought up, etc. Apart from these considerations, the behavior ratings were obviously sufficiently reliable to yield significant and consistent experimental effects.

The Liebert reference to the analysis of the 22 raters is misplaced since it was not intended to evaluate 'inter-rater agreement" although reliability does indirectly enter into this comparison. The principal objective of this analysis was to reduce the influence of any single rater upon the experimental findings by only using the rater's mean score for the control subjects and for the aggressive TV subjects (for those raters who had observed subjects in both groups). The fact that a significant difference at the .05 level was obtained with only 22 degrees of freedom as compared to over 600 degrees of freedom when we used individual subject scores as our unit of comparison did seem to us to be impressive (and it still does).

(2) The potential problem posed by possible differences in preference for aggressive and nonaggressive programs was a serious one. From the beginning of the study, we took a number of steps to minimize the possible influence of this factor. First, in most of the participating institutions, we used volunteers. Secondly, we paid the boys to participate (and, in addition, provided them with new large screen television sets). Also, the boys could select the particular programs they wished to watch, provided they were on the assigned diet, and, in addition, there were many more control than aggressive television programs from which they could choose. Finally, in the case of one objection that some of the boys in the California Control group raised, namely, their desire to see *Batman*, we resolved it in favor of minimizing frustration. In terms of the design, it was essential that not all, but the *preponderance* of programs witnessed, be either aggressive or nonaggressive, depending upon whether the boy was in the aggressive television or control condition. As the data indicate, this objective was achieved.

In addition to implementing procedures designed to minimize frustration, we also carried out a number of statistical analyses aimed at evaluating the hypothesis that witnessing the control programs was frustrating and possible responsible for the greater aggression of the control as compared to the aggressive television group. If one assumes that frustration was operating for the controls, it is reasonable to infer that boys who had watched a great deal of television should have been more frustrated by being placed on the control diet than boys who had watched relatively little television and hence should show a stronger experimental effect. If this had proved to be the case, it would have been evidence for the frustration assumption. The fact that this was not the case is evidence against frustration effects—not conclusive evidence by any means, just one type of data to be used in conjunction with other pertinent data. Other pertinent data are the findings that two of the control groups in the four institutions in which significant differences in behavior ratings were obtained between the control and aggressive television treatment groups showed significant *increments* in their mean liking of nonaggressive programs after the six-week period (the initial mean rating for all four control groups was already high). In contrast, the private school controls are less enamoured with nonaggressive programs after exposure to the control diet and yet there was no evidence of an increase in the behavior rating of aggression in these private school controls. Also relevant are the *positive* correlations obtained between liking of the sample of aggressive programs and liking of the nonaggressive programs used in the pre and post assessment.

We finally turn to the reaction to the programs actually watched by the boys. The "dislike" ratings were combined in the verbal description since the percentages were so small in each of the three subcategories. The category "liked a little" was omitted from the verbal description in order to highlight the predominantly positive attitude toward both the control and aggressive television programs that were witnessed. As for the "15 percent of the data—omitted entirely," the missing 15 percent can actually be recovered from the data given. If one adds the 70 percent of the control programs "liked very much" and "fairly much" to the 15 percent "dislike" ratings and subtract the sum from 100 percent, we obtain the missing 15 percent which is precisely the percentage in the "liked a little" category. The same percentage of 15 percent is obtained for the aggressive programs that were "liked a little." Conceding that this arithmetic operation may not have been obvious from the text, the reviewers might have sent for a copy of the table listed as #31 (Reactions of the Control Group to Non-Aggressive TV Programs and the Experimental Group to Aggressive TV Programs) in the supplementary materials which are listed in the text as available from the authors on request. Included in the information on this table are the following data:

	Liked it very much	Liked it fairly much	Liked it a little	Disliked it a little	Disliked it fairly much	Disliked very much
Control	47%	23%	15%	5%	3%	7%
Agg. TV	51%	26%	15%	3%	2%	3%

Also included in this table and described in the text is an analysis of affective reactions, the aggressive programs tending to elicit more *negative affective* responses and the control programs eliciting more *positive affective* responses.

While no single analysis is in itself conclusive, the combination of all of these observations indicates that the hypothesis of differential resentment or frustration is an improbable one.

(3) There are two principal issues subsumed within the overall question of the influence of possible response bias of the raters. (a) What expectation did the raters entertain concerning various aspects of the study? (b) Can one reasonably account for the findings by hypothesizing response bias on the part of the raters?

The question of possible rater bias effects is clearly an important one, particularly in a study where the raters were familiar with the experimental treatment (although not necessarily of its import). With regard to the issue of rater expectations as contrasted with the related question of rater influence, this issue cannot be fully resolved by interview since many expectations remain implicit and unverbalized. What we can say concerning the raters' expectations is that only one rater at any time during the course of the study overtly expressed *concern* about the experimental assignment. This rater, a cottage mother in one of the boys' homes, in the first week of the study wanted to switch two boys who were in the aggressive programs group to the control television programs because she felt that the aggressive television programs were too exciting and might get them "to act up" (the transfer was not permitted).

The experimental design was not discussed with the raters prior to completion of the study and consequently we did not assess their initial expectations. (The institution authorities, with whom we did reveal the full purpose of the study, either believed aggressive television to be irrelevant to aggressive behavior or with one exception thought that it stimulated aggression in some children.)

The objections raised to the frustration assumption also apply to the Liebert et al. conjecture concerning the rater's possible perception of the effects of removal of the boys from their favorite program. The question of the rater's possible insight into the purpose of the study is more difficult to assess. Of the 12 raters interviewed following completion of the study, only two conveyed insight into the project (both raters were at private schools). It was not that the other raters were unaware that they were recording aggressive incidents. They believed (according to their verbal statements) that the study was concerned with the relationship between the boys' personalities (as measured by the behavior ratings and the questionnaires given to the boys) and their ratings of the television programs on their "diet."

However, the critical question is not the expectancies which the raters may or may not have entertained but whether the results can be accounted for by invoking rater bias. A minor objection to the rater bias hypothesis is that it does not account for the absence of significant effects in the private schools. A major objection is the differential experimental effect found as a function of personality variations among the boys. While it is difficult to believe that the rater bias worked selectively with boys initially more aggressive and undercontrolled, it "stretches our credulity" to believe that the raters were able to discriminate low fantasy aggression from high fantasy boys and selectively attribute more aggression to the low fantasy aggression boys in the control group. A second major objection is that there were other data gathered besides the behavior ratings. The results of the pre-post measures based on the responses of the *boys* were *consistent* with the findings based on the behavior ratings. The differences between the aggressive television and control groups found on the questionnaire and sociometric data, while not as impressive as the effects obtained on the behavior rating measures, were nevertheless statistically significant for various personality subgroups, in most instances the same personality subgroups that showed the strongest experimental effects on the behavior ratings.

In view of these several considerations, it would be very difficult to account for the obtained data on the assumption of rater bias and, on these grounds, the rater bias hypothesis is rendered highly unlikely.

(4) It is difficult to give the *Batman* hypothesis the same serious consideration as the first three objections that were raised. In satisfying the expressed wishes of some of the controls to see *Batman,* we did treat them differently than the experimental group who had made no comparable request. We thereby reduced possible frustration and, if anything, reinforced such behaviors as requesting permission, seeking appropriate recourse rather than, as Liebert et al. contend, "grumbling, breaking rules, refusing tasks," etc. However, rather than engage in conjectures as to what was reinforced and what possible implications, remote or otherwise, this action may have had, an examination of the experimental results provides an adequate basis for the rejection of the *Batman* hypothesis as a possible source of bias. Allusion has already been made to the interaction of the experimental effects with personality variables. While one might stretch the Batman hypothesis to account for these interactions, it would have greater difficulty in accounting for the significant *decline* in aggression (as assessed by the behavior ratings) that took place over the six-week experimental period in the group assigned to the aggressive television diet. More germane and a fundamental objection to the Batman hypothesis is the fact that *both* New York institutions, where the Batman exception was *not* made, showed significant experimental effects. At the one boys' home in which the difference in behavior ratings of peer aggression between the Controls and Aggressive television group was significant at only the .10 level, the difference in authority aggression was significant at the .001 level. We think it reasonable to assert that a significant experimental effect took place in this boys' home as well as in the other boys' homes. Inasmuch as the experimental effect in the two boys' homes in which the controls did not watch *Batman* was comparable to the effect obtained in the two boys' homes in which the controls were permitted to watch *Batman,* the deviation from experimental orthodoxy entailed in granting the strong request to watch *Batman* is irrelevant to the experimental outcome.

All of the supposed major "weaknesses" cited by Liebert et al., assume minor proportions or disappear with a careful reading of the text. One would never gather from this review that the *Television and Aggression* volume reports a set of theoretically coherent and empirically consistent and significant findings. Reasonable inferences are possible from the study even though the findings happen to run counter to the evidence provided by studies using very different procedures and subject populations. Parenthetically, there are data provided by other studies which are consistent with the *Television and Aggression* findings. Aside from the supporting results of other experiments in which we have been involved, research stemming from the laboratory of Professor Jerome Singer among others has also provided evidence of catharsis-like effects resulting from the observation of aggressive television content. This is not to say that the catharsis hypothesis is more valid than the modeling or stimulation hypotheses. Rather, the appropriate scientific task seems to us to be one of developing a theoretical system that can incorporate individual differences, situational contexts and variations in television content—in brief, that can handle the diverse findings in this area rather than deny them.

Catharsis of Aggression Among Institutionalized Boys: Further Comments

Robert M. Liebert, Emily S. Davidson, and Michael P. Sobol
State University of New York at Stony Brook

In replying to Liebert, Sobol, and Davidson (1971), the authors of *Television and Aggression* have not questioned the importance of the several methodological issues which were raised. Rather, Feshbach and Singer (1971b) have asserted that these problems are not serious in the investigation which they have reported. In doing so they have presented some information not available to us from their published report, clarified certain aspects of their theoretical and methodological orientation, and brought to light a number of misunderstandings in the interchange which has transpired thus far. Therefore, a few further comments appear to be in order. We shall not, however, reiterate either those of our earlier comments which have not been challenged or those in which the issue seems to be simple disagreement rather than misunderstanding.

Feshbach and Singer (1971b) take exception to our assertion that the many positive correlations obtained by other investigators (between viewing aggressive television and aggressive behavior) are at odds with the catharsis hypothesis. They imply that the positive correlations are obtained not because (1) exposure to such content leads to aggressive behavior, but rather because (2) aggressive boys are particularly drawn to aggressive television content. While the correlational studies themselves cannot be used to prove conclusively that observing televised aggression causes aggression, Chaffee and McLeod (1971) have pointed out that these alternatives [(1) and (2)] can be compared. After reviewing several other lines of pertinent evidence which favor the former over the latter, these investigators note:

> Other data render H_2 [''aggressiveness causes adolescents to watch aggressive television programs''] even less plausible to the extent that it implies a high degree of purposefulness in adolescent program selection. Lyle and Hoffman (1971) found that TV is primarily associated with entertainment and relaxation (not anger or hurt feelings) among adolescents. A sample of boys in early adolescence said they often watch television to kill time, or because they ''just came on'' (Friedman, 1971). And the low correlations we find here between previous and present violence viewing would suggest that specific behavior is not often a fixed ''personality trait'' that remains associated with other more basic traits; Lefkowitz, *et al.*, (1971) report significant longitudinal test-retest correlations for aggressiveness *but nor for viewing preferences* (Chaffee and McLeod, 1971, p. 17, italics added).

While we regret failing to present a discussion of the ''reverse-cause'' hypothesis for brevity's sake, it does not seem to account for the abundant data to which we referred.

Although Feshbach and Singer (1971b) have conceded that their non-parametric analyses overlooked curvilinear trends,[2] they defend these statistics as being ''more conservative and more *justified*'' than the parametric alternatives. But, in accounting for a seemingly clear difference which might appear to be at odds with their conclusions, Feshbach and Singer note: ''In the private school high aggression group the aggressive TV group showed a substantially higher overall median [2.24 vs. 1.30 for aggression toward peers] than the controls. In this instance the median difference tends to *exaggerate* the actual difference between the two groups. (The means for the aggressive TV and control groups are 2.39 and 2.38.)'' [1971a, p. 85, italics added]. We are not challenging the qualification which appears in the foregoing statement. Its import is in illustrating that non-parametric statistics are not always more conservative, or more justified, than parametric ones. (Parametric statistics are, however, often favored in order to discriminate the presence or absence of overall effects and to determine whether interactions are actually reliable before employing individual comparisons.)

The authors of *Television and Aggression* describe our statement that the actual instrument used was never checked for reliability as an ''unfortunate exaggeration.'' Specifically, they have replied that only one of the 26 items on the behavior rating scale employed at the Fernald School was later modified. Given this information, we agree completely that the instrument used at that institution and in the actual study reported by Feshbach and Singer did not differ materially in content. One might have thought that somewhat more extensive modifications had been made as a result of the check described, from this statement in the original text:

"The behavior rating scale was pretested on boys at the Fernald School at UCLA and, on the basis of the pretest, was revised" (1971a, p. 56). However, in a more important sense, it remains true that the instrument was not tested for reliability.

For example, it is possible that the classroom teacher and teaching supervisor who were involved in the Fernald School check would have agreed substantially on the relative (or rank order) aggressiveness of the various boys under their supervision regardless of what transpired on a given day. To the extent that this might be the case, the check performed does not necessarily provide any information about the reliability of the instrument itself. Such information only becomes available when two or more raters (preferably unfamiliar with the subjects prior to the ratings) independently agree on the actual events that occur in the presence of both of them. That one can be misled by results obtained in the absence of such a check has been clearly demonstrated by Reid (1970), to whose work we have alluded in our earlier review.

An intriguing argument, advanced on two different occasions in the Feshbach and Singer reply, is that the reliability issue can be resolved by the fact that the rating measure yielded some significant effects. Clearly, no more than a chance number of significant effects could have resulted if the ratings were completely random. However, if this argument obviated the need for reliability checks, researchers who obtain significant effects would almost never be required to report independent reliability data. But such checks are required (cf. Anastasi, 1968). Among other problems, a measure with poor or unassessed inter-rater reliability can produce effects that are due to only a subsample of the raters. In turn, such effects would be particularly troublesome regarding the problem of rater bias (Cronbach, 1960; Reid, 1970) if effects are obtained.[3]

Regarding the bias issue, our initial review explicitly noted that raters were unlikely to have entertained the catharsis hypothesis. That possibility did not therefore require rebuttal. What was not addressed in the reply is our suggestion that many raters would be drawn to the following reasoning: "Most boys liked aggressive programs better than nonaggressive ones before the experiment began [in fact, this initial difference was statistically significant (p. 109] and many watched such shows regularly. When deprived of something which they like and have previously enjoyed, they will become more irritable and thus more likely to be more aggressive and rowdy." As was pointed out previously, the raters can entertain this view (and it can influence their ratings) without their assuming either that such an hypothesis was held by the experimenters or that detecting the phenomenon was a purpose of the research.

Countering our suggestion that rater bias might have accounted for the obtained effects, Feshbach and Singer argue that the bias would not favor attributing a stronger effect to those who were aggressive and undercontrolled initially. It seems to us that the bias would tend to run in *just this direction*. Initially aggressive boys, if deprived of shows that had previously been their favorites, might be the very ones most likely (from the raters' point of view, at least) to react with increased aggression. In the same vein, they write: "it 'stretches our credulity' to believe that the raters were able to discriminate low fantasy aggression from high fantasy boys and selectively attribute more aggression to the low fantasy aggression boys in the control group" (1971b, p. 9). Since the median difference between high and low fantasy boys in peer aggression (Table 13, p. 92) is descriptively smaller in the control than in the aggressive diet group, we presume they must be claiming differences in the magnitude of their experimental effect as a function of high and low fantasy. The claim is not supported. There is a difference in the *p* levels associated with the reported effects but, especially since both show the same general tendency, it would be inappropriate to speak of a differential effect. Doing so would be equivalent to seizing upon the significant ($p < .05$) linear decrease among those exposed to the *non*aggressive diet in the private schools (versus no linear trend in the aggressive program group) and arguing for a modeling effect. Feshbach and Singer (1971a, p. 78) have themselves pointed out the need for comparing the differences statistically before advancing any such argument.[4]

The remaining responses made to the rater bias contamination deal with measures other than rated behavior. To respond adequately to the rater bias question through these other data, two arguments must be invoked. First, it must be agreed that the raters could not have influenced these measures. Second, it must be shown that the results obtained on nonrated measures might have effectively stood on their own in terms of the conclusions drawn. The first point was never challenged.

However, a careful reading of the text reveals that the second point is not supported compellingly. Rather than belabor the details, we would note the investigators' own description of these supplementary data:

> Perhaps the most salient aspect of these data is that the differences are not as impressive as those obtained in the analysis of aggression toward peers and toward authority. Where significant experimental effects were found they usually held for personality subgroups rather than for the whole sample. However, except on the measure of change in aggressive fantasy, the differences were consistent with the behavior aggression data, the control subjects tending to develop more favorable attitudes and values toward aggression than the boys exposed to aggressive content on television. *However, this statement must be carefully qualified.*[5] *The results differ for each measure; some experimental effects hold only for the private schools or for the boys' homes or only for particular personality dispositions* (1971a, p. 97, italics added).

We do not feel that these supplementary data supplant the rater bias account as one possible competing hypothesis for the behavioral effects.

In responding to the confound which Feshbach and Singer have dubbed the "Batman hypothesis," it is implied that the experimental effect can stand on the basis of those schools in which *Batman* was not seen. There are two bases for questioning this argument. First, of the eight individual comparisons involving control groups which are not confounded in this way, only two reach significance.[6] Second, in only one ("B") of the schools in which an experimental effect is claimed did the investigators detect a significant difference on both peer and authority aggression; "B" is one of the schools in which Feshbach and Singer did accede when, in their words, control group boys at these institutions "*objected very strongly* because *Batman* was not on their list . . ." (1971a, p. 55, italics added).

It is also suggested, in Feshbach and Singer's reply to the Batman hypothesis, that we must account for the decline shown by the aggressive TV group in the boys' homes. We would offer their own argument that "an increment or decrement in a particular group is not very revealing" (1971a, p. 78) together with the fact that the first and seventh week medians for this group (1.43 and 1.36, respectively) were virtually identical.

Since we did not wish to hazard assumptions regarding the measure of program liking (assumptions required *before* performing the arithmetic operations which Feshbach and Singer have detailed for us), we appreciate full publication of the liking data for aggressive and control diets. Feshbach and Singer also observe that: ". . .the reviewers might have sent for a copy of the table listed as #31 (Reactions of the Control Group to Non-Aggressive TV Programs and the Experimental Group to Aggressive TV Programs) in the supplementary materials which are listed in the text as available from the authors on request" (1971b, p. 7). To avoid any unfortunate impressions, it should be noted that *Television and Aggression* lists only 30 such items (1-30), none of which bears the title shown, in its chapter "Note on Supplementary Materials" (pp. 172-74).

Feshbach and Singer did not supplement the tabular information on liking with a statistical analysis, but Chaffee and McLeod (1971) have reported such a test since our review was written. They found that significantly more of the control than aggressive diet programs (15 percent *vs.* 8 percent) were rated as disliked.[7] Although our concerns extend beyond the frustration issue, Chaffee and McLeod found it alone sufficient to discount the interpretation advanced by Feshbach and Singer, suggesting:

> A reasonable interpretation might be that "high violence viewing" is nearer to the normal pattern of TV use for boys in early adolescence (Friedman, 1971; Lyle and Hoffman, 1971; McIntyre and Teevan, 1971), so that the "low violence viewing" condition constituted a more artificially instituted pattern of deprivation; the boys reacted to this constraint, rather than to what they watched on TV (1971, pp. 26-27).

We have suggested previously that the personality data are neither strong enough, nor consistent enough, to dismiss the rater bias problem. However, it should be noted the disruption-frustration hypothesis might well account for the few individual comparisons that are significant. In this vein, it is of further interest that the treatment referred to as the control group (a convention we have accepted for ease of communication) experienced at least as much disruption of their normal viewing habits as those exposed to the predominantly aggressive programs. We hope to conduct a constructive replication in which, beyond attention to the other problems in design, a third, true control, group is observed without any manipulation of television viewing. This would be especially desirable since the data reported in *Television and Aggression* which could be pertinent to the baseline problem confound a pre-experimental observation period with the first week of treatment.

Feshbach and Singer (1971b) have observed that we did not accept the findings reported in *Television and Aggression* as a "theoretical challenge and opportunity." Given the many alternative explanations which can account, individually or in combination, for their data, we continue to feel that it was (and still would be) injudicious to build an elaborate theoretical structure (or speak to the issue of broadcasting policy) without a more compelling foundation.

FOOTNOTES

1. The Lefkowitz, Eron, Walder, and Heusmann (1971) study provides even clearer evidence than is suggested by Chaffee and McLeod's comments. Whereas the association between aggressiveness in the third grade and violence viewing at age 19 accounts for about one-tenth of one percent of the variance in their data, the relationship between early violence viewing and later aggression accounts for roughly ten percent of the variance, and is significant. The conclusion which they draw from their data, that violence viewing is causally implicated in aggression, has been accepted by critics who have first imposed the most stringent methodological standards in reanalysis (Campbell, 1971; Neale, 1971).
2. It is of passing interest that nonparametric techniques are available for detecting curvilinear trends (Morrison, 1971).
3. Reid (1970) specifically notes that, in the absence of high reliability, ". . . faulty research hypotheses might find *more* than their share of statistical support" (p. 1149, italics added).
4. The issue of a criterion for significant effects is a complicated one. It should be noted that $p < .10$ effects are sometimes labeled as NS (nonsignificant) when they show instigating or "modeling" effects, a decision which is exemplified in the physical aggression toward peers comparisons for boys who were initially aggressive in the private schools (Table 11, p. 86 and the text on p. 87). Thus the customary significance criterion of $p < .05$ is consistently employed for those findings which we have had occasion to review.

5. Of the 36 comparisons presented in Tables 14-16, only six reach statistical significance. In the remaining table in the chapter, in which the exception alluded to (aggression in fantasy)is shown, three of 12 comparisons are significant.
6. In evaluating the results reported in *Television and Aggression*, it is important to note that the division into private schools and boys' homes for purposes of analysis is *post hoc*, since the authors specifically note (p. 74) that no differential predictions were advanced regarding type of institution.
7. Chaffee, in personal communication, has indicated that by his computations the significance level is $p < .001$.

REFERENCES

Anastasi, A. *Psychological testing* (Third Ed.). New York: Macmillan, 1968.

Campbell, D. T. Open letter to Eli A. Rubinstein, 1971. (In support of a document by D. A. Kenny: Threats to the internal validity of cross-lagged panel inference, as related to "Television violence and child aggression: a followup study.")

Chaffee, S. H., and McLeod, J. M. Adolescents, parents, and television violence. Paper presented at American Psychological Association annual meeting, Washington, D.C., September 1971.

Cronbach, L. J. *Essentials of psychological testing* (Second Ed.). New York: Harper & Row, 1960.

Feshbach, S., and Singer, R. *Television and aggression*. San Francisco, California: Jossey-Bass, Inc., 1971 (a).

Feshbach, S., and Singer, R. D. *Television and aggression:* a reply to Liebert, Sobol, and Davidson. Report submitted to National Institute of Mental Health, 1971 (b). (Preceding document)

Friedman, H. L. and Johnson, R. Mass media use and aggression: a pilot study. In *Television and social behavior*, Vol. 3. Washington, D.C.: U. S. Government Printing Office, 1971.

Lefkowitz, M. M., Eron, L. D., Walder, L. C., and Huesmann, L. R. Television violence and child aggression: A followup study. In *Television and social behavior*, Vol. 3.

Liebert, R. M., Sobol, M. P., and Davidson, E. S. Catharsis of aggression among institutionalized boys: Fact or artifact? Report submitted to National Institute of Mental Health, 1971. (In preceding pages)

Lyle, J., and Hoffman, H. R. Children's use of television and other media. In *Television and social behavior*, Vol. 4.

McIntyre, J. J., and Teevan, J. J. Television and deviant behavior. In *Television and social behavior*, Vol. 3.

Morrison, H. W. Distribution-free tests of non-linear trend. Unpublished memorandum, 1971 (Based upon standard statistics texts and available upon request).

Neale, J. M. Comment on: Television violence and child aggression: A followup study. In *Television and social behavior,* Vol. 3.

Reid, J. B. Reliability assessment of observation data: a possible methodological problem. *Child Development,* 1970, **41** 1143-50.

Television and Aggression: Some Reactions to the Liebert, Sobol and Davidson Review and Response

Seymour Feshbach

University of California, Los Angeles

and

Robert D. Singer

University of California, Riverside

The Liebert, Davidson, and Sobol comments on our response are more tempered than those made in the initial critique and, for the most part, leave unchallenged the statements and data cited in our rejoinder. Some additional points are raised and some technical issues debated. We shall limit this second response to those issues that are germane to the interpretation of the principal findings and to a clarification of misunderstandings that still remain.

For example, possible misunderstandings might arise from the following statement (made in connection with their discussion of the relative merits of parametric versus nonparametric statistics): "But in accounting for a seemingly clear difference which might appear to be at odds with their conclusions, Feshbach and Singer note . . ." First, as the text indicates, the particular difference cited was not statistically significant, and our allusion to the mean difference as made in an effort to help clarify why this was the case rather than in an effort to account for a supposedly deviant finding. More important, a significant effect in a private school would not have been "at odds" with either the hypotheses proposed at the initiation of the study or with the theoretical interpretation given to the findings. We expected evidence of an aggression-stimulating effect resulting from exposure to the aggressive television diet in certain personality subgroups and searched the data for such effects. We believe that most readers of the book will agree that particular attention was given to any differences indicating heightened aggression in the aggressive television group.

As to the parametric nonparametric issue, it bears little relevance to the experimental outcome and, at the risk of repetition, the skewedness of the distributions governed the decision to employ non-parametric procedures in the trend analyses. The import of their reference to the unpublished Morrison (1971) memorandum, which apparently was written about the time our book appeared, is unclear, especially since it does not seem to be concerned with a nonparametric statistic for determining the reliability of a difference between two curvilinear trends, a statistic which would be critical to any nonlinear comparisons between the Aggressive and Control TV groups.

We did make an error which we wish to acknowledge in stating that item 31 was included in the list of supplementary materials enumerated in the book. It was included in the manuscript sent to the publisher and in the packet sent to the various individuals who have requested the supplementary items. However, it did not appear in the printed list. The reference to the Chaffee and McLeod statistical analysis of these data is rather puzzling. The reason why we did not supplement the tabular information on liking with a statistical analysis is that the data reported do not lend themselves to the usual test of statistical reliability. The figures given describe proportions of total programs without regard to subject variability or differential contribution to these proportions. Consequently, the meaning of a statistical test is ambiguous. We do not doubt that there may be a reliable difference in preference for the aggressive diet as compared to the control programs. Nonetheless, the difference is small, the great majority of both types of programs received positive ratings, and given such evidence as the increase in preference for nonaggressive programs in the Boys Homes controls and other data alluded to in our previous response, we believe it reasonable to conclude that the experimental results are not due to possible frustrating effects of the control diet.

The other possible sources of methodological bias cited by Liebert, Davidson, and Sobol are essentially low probability alternatives that can be rejected by an examination of the actual findings. The relationship between the reliability of the behavior ratings and possible rater bias is very tenuous, particularly in view of the correlations of the behavior ratings with self and peer ratings. Apart from the informal observation that any rater bias is more likely to work in the direction of attributing greater aggression to subjects in the aggressive television group, the changes in aggression on the questionnaire and sociometric measures, even though less powerful and more qualified than the finding for the rating measures, provide evidence of statistically reliable experimental effects which are independent of possible rater bias. As for the *Batman* hypothesis, it is both highly conjectural and, despite the Liebert et al. count of both private school and Boys Homes control comparisons, inconsistent with the observation of significant experimental effects in the two Boys Homes in which the *Batman* issue did not arise. We wonder what Liebert et al. might have written concerning the greater aggression in the Controls if we had denied the request of the Control subjects who had asked for permission to watch their favorite program, *Batman*.

The relevance of a methodological problem is also dependent upon the type of inference to be made from particular experimental findings. Thus, there may be unknown factors correlated with the experimental treatment which are responsible for the experimental differences that were found. Apart from field investigations, there have been very few laboratory studies of media effects with predominantly poor, nonacademically oriented preadolescents and adolescents to which one might refer for additional insight into the nature of the mediating processes. The results for the Boys Homes populations clearly need to be replicated with other samples and in other settings before one can be confident of the aggression-moderating effects that aggressive television content may have for certain types of children.

We are more confident, however, in concluding that the type of aggressive programs which these boys watched had no significant aggression-stimulating influence. Liebert, Davidson, and Sobol do not speak to the absence of aggression-stimulating effects on the self-report measures, the sociometric indices, and most of the behavior ratings (a few of the many analyses of behavior ratings suggesting possible small effects in some private school subgroups). If there were aggression-stimulating effects pursuant to the observation of the aggressive television diet, they appear to be of little pragmatic consequence.

This is not to exclude the possibility that television shaping effects may already have taken place by the time boys reach preadolescence and that young children may show very different effects than older children when exposed to aggressive content in television. Also, the effects of aggressive content in television may vary with such factors as sex, prior exposure to television, socioeconomic level, cognitive resources, emotional state, and aggressive behavioral history of the audience and such structural characteristics of the program as the outcome of the aggressive act, the reality of the dramatic content, the tension inducing-relaxing properties of the drama, and so on. It would be very surprising, indeed, if the effects of aggressive content were so simple and uniform as to render unnecessary a theoretical structure which could account for diversity in response. And, in like vein, we would agree with Liebert et al. that one needs a compelling foundation before speaking to the issue of broadcast policy.

☆ U. S. GOVERNMENT PRINTING OFFICE : 1972 720-718/635

DHEW Publication No. HSM 72-9060
Printed 1972

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Health Services and Mental Health Administration

National Institute of Mental Health